

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problems Mailbox.**

**This Page Blank (uspto)**

(19) 日本国特許庁(JP)

## (12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平 10-117789

(43) 公開日 平成10年(1998)5月12日

(51) Int. Cl. <sup>6</sup>	識別記号	FI
C 1 2 N 15/09	Z N A	C 1 2 N 15/00 Z N A A
A 0 1 K 67/027		A 0 1 K 67/027
A 6 1 K 35/12		A 6 1 K 35/12
38/46	A D S	48/00 A A M
48/00	A A M	C 1 2 N 9/64 Z
審査請求 未請求 請求項の数 27	FD 外国語出願	(全 1 4 2 頁最終頁に続く)

(21) 出願番号 特願平9-281032

(22) 出願日 平成9年(1997)9月5日

(31) 優先権主張番号 60/025436

(32) 優先日 1996年9月6日

(33) 優先権主張国 米国 (U S)

(31) 優先権主張番号 60/027873

(32) 優先日 1996年10月25日

(33) 優先権主張国 米国 (U S)

(31) 優先権主張番号 60/032875

(32) 優先日 1996年12月13日

(33) 優先権主張国 米国 (U S)

(71) 出願人 595047190  
 スミスクライン・ビーチャム・パブリック  
 ・リミテッド・カンパニー  
 SmithKline Beecham  
 p. l. c.  
 イギリス国ミドルセックス・ディーダブリ  
 ュ8・9イービー、ブレンフォード、ニュー  
 ・ホライズンズ・コート (番地の表示なし  
 )

(74) 代理人 弁理士 青山 葆 (外1名)

最終頁に続く

(54) 【発明の名称】 ヒトセリンプロテアーゼ

## (57) 【要約】

【課題】 P S P 1 に対応するヌクレオチドおよびアミノ酸配列の提供、P S P - 1 に結合するP S P - 1 のモジュレーターおよび/またはP S P 1 タンパク質分解活性のモジュレーター、そのモジュレーターの同定方法、ならびにかかる方法に有用な物質が必要とされている。

【解決手段】 ヒトセリンプロテアーゼP S P 1 をコードする単離核酸、当該核酸より入手できるタンパク質、当該核酸で形質転換した組換え宿主細胞、P S P 1 多型性に特異的なオリゴヌクレオチドおよびプライマー対、ならびに当該タンパク質および当該核酸配列の使用を開示する。

## 【特許請求の範囲】

【請求項1】 生物学的に活性なPSP1ポリペプチドをコードする単離ポリヌクレオチド。

【請求項2】 (a) ヌクレオチド603から1979の配列番号：24に示すヌクレオチド配列を有するPSP1-1をコードするポリヌクレオチド；

(b) ヌクレオチド603から1979の配列番号：23に示すヌクレオチド配列を有するPSP1-2をコードするポリヌクレオチド；

(c) ヌクレオチド603から1736の配列番号：26に示すヌクレオチド配列を有するPSP1-3をコードするポリヌクレオチド；

(d) ヌクレオチド603から1913の配列番号：28に示すヌクレオチド配列を有するPSP1-4をコードするポリヌクレオチド；および

( ) D87257 (1325T) タンパク質をコードするポリヌクレオチド；からなる群から選択される単離ポリヌクレオチド。

【請求項3】 配列番号：24；配列番号：23；配列番号：26；または配列番号：28に実質的に類似する単離ポリヌクレオチド。

【請求項4】 ヌクレオチド672および1435が独立してCおよびTから選択される請求項2または3記載の単離ポリヌクレオチド。

【請求項5】 配列番号：23、24、26、28もしくは30またはヌクレオチド1325がTである配列番号：17に示すヌクレオチド配列を有する単離ポリヌクレオチド。

【請求項6】 請求項1から5記載のいずれか1つのポリヌクレオチドによりコードされる機能性ポリペプチド。

【請求項7】 配列番号：25または30に示すアミノ酸配列を有するPSP1-1；配列番号：8に示すアミノ酸配列を有するPSP1-2；配列番号：27に示すアミノ酸配列を有するPSP1-3；配列番号：29に示すアミノ酸配列を有するPSP1-4；または配列番号：18に示すアミノ酸配列（アミノ酸残基213がvalである）を有するD87257 (1325T) タンパク質；である請求項6記載の機能性ポリペプチド。

【請求項8】 DNAまたはRNAである請求項1ないし5記載のいずれか1つのポリヌクレオチド。

【請求項9】 請求項8に記載のDNAを有してなるベクター。

【請求項10】 請求項9に記載のベクターを有してなる組換え宿主細胞。

【請求項11】 タンパク質の発現を促進する条件下、請求項10に記載の組換え宿主細胞を培養し、発現したタンパク質を回収することからなる、本質的に純粋なPSP1タンパク質またはD87257 (1325T) タンパク質の製造方法。

【請求項12】 請求項11に記載の方法により製造したPSP1またはD87257 (1325T)。

【請求項13】 請求項1ないし5に記載のいずれか1つのポリヌクレオチドまたはD87258との結合能を有する配列を有してなるアンチセンスオリゴヌクレオチド。

【請求項14】 請求項6またはD87258のポリペプチドのモジュレーター。

【請求項15】 PSP1またはD87258活性を調節する物質の存在について培地を分析する方法であって、

(i) (a) 請求項8に記載のPSP1ポリペプチドもしくはD87258タンパク質、またはその機能性誘導体および細胞性結合パートナーまたはその合成類似体を得；

(b) PSP1またはD87258タンパク質／細胞性結合パートナー複合体を形成しうる条件下で、PSP1またはD87258活性を調節すると考えられる試験物質と一緒にインキュベートし；

(c) その複合体、遊離PSP1またはD87258タンパク質または遊離細胞性結合パートナーの存在を分析し；および

(d) 対照と比較して該物質の効果を決定する；工程からなる、PSP1またはD87258の細胞性結合パートナーへの結合に影響を及ぼすこと；

(i i) (a) 請求項8に記載のPSP1ポリペプチドもしくはD87258、またはその機能性誘導体および細胞性基質またはその合成類似体を得；

(b) PSP1またはD87258酵素／基質複合体を形成し、その後の基質の切断を可能とする条件下で、PSP1またはD87258活性を阻害すると推測される試験物質と一緒にインキュベートし；

(c) タンパク質分解的に切断した基質の存在を分析し；および

(d) 対照と比較して該物質の効果を決定する；工程からなる、細胞性基質に対するタンパク質分解活性を阻害すること；

(i i i) (a) 請求項8に記載の標識化PSP1ポリペプチドもしくはD87258、またはその機能性誘導体を得；

(b) 固体支持体に結合したモジュレーター候補物質を得；

(c) PSP1またはD87258タンパク質／モジュレーター候補物質複合体を形成しうる条件下で、標識化PSP1またはD87258タンパク質と固体支持体に結合したモジュレーター候補物質の混合物をインキュベートし；

(d) 固体支持体を遊離可溶性標識化PSP1またはD87258タンパク質から分離し；

(e) 固体支持体に結合した標識化タンパク質の存在を

分析し；

(f) 標識化PSP1またはD87258タンパク質と複合体形成した固体支持体を単離し；および

(g) モデュレーター候補物質を同定する；工程からなる、PSP1またはD87258タンパク質に直接的に結合させること；を特徴とする分析方法。

【請求項16】 請求項15に記載の方法により同定されるPSP1またはD87258タンパク質調節化合物。

【請求項17】 PSP1またはD87258活性を調節する必要性を有する患者の治療薬の製造における、請求項16に記載の調節化合物の使用。

【請求項18】 請求項16に記載の調節化合物および医薬上許容される担体を有してなる医薬組成物。

【請求項19】 (a) 個体からポリヌクレオチドサンプルを単離し；

(b) ポリヌクレオチドサンプルおよび、配列番号：23、24、26、28もしくは30に示すヌクレオチド配列を有するPSP1をコードするポリヌクレオチド、または配列番号：18に示すD87258をコードするポリヌクレオチドを分析し；および

(c) 任意の差異がPSP1またはD87258配列における突然変異を示す、ポリヌクレオチドサンプルおよびPSP1またはD87258ポリヌクレオチド間の差異を比較する；ことからなる、PSP1またはD87258タンパク質欠乏に付随する状態の診断方法。

【請求項20】 PSP1またはD87258タンパク質機能不全に関与する状態の治療方法であって、

(i) (a) PSP1またはD87258タンパク質機能が欠損している患者から細胞を単離し；

(b) 請求項1ないし7に記載のいずれか1つのポリヌクレオチドまたは配列番号18に示すD87258をコードするポリヌクレオチドを、PSP1またはD87258タンパク質を発現する細胞にトランスフェクトすることにより細胞を改変させ；および

(c) 細胞を患者に導入して戻し、状態を緩和する；工程からなり；または

(i i) 請求項1ないし5に記載のいずれか1つのポリヌクレオチドまたは配列番号18に示すD87258をコードするポリヌクレオチドを、PSP1またはD87258タンパク質機能が欠損している患者に投与する方法であって、PSP1またはD87258タンパク質を発現させ、状態を緩和することからなる治療方法。

【請求項21】 PSP1、D87258またはその免疫抗原と免疫反応する抗体。

【請求項22】 トランスジェニックなヒト以外の動物の任意の細胞で、請求項5に記載のDNAまたは配列番号：18に示すD87258をコードするポリヌクレオチドの発現能を有するトランスジェニックなヒト以外の動物。

【請求項23】 患者、好ましくはアルツハイマー病に対する神経変性素因を有する患者由来のサンプル中のPSP1またはD87258多型性を検出することからなる、患者における神経変性についての遺伝的素因の決定方法。

【請求項24】 多型性がPSP1のヌクレオチド672、PSP1のヌクレオチド1435、またはD87258のヌクレオチド1325で検出される、請求項23記載の方法。

【請求項25】 多型性がポリメラーゼ連鎖反応により検出され、好ましくはポリメラーゼ連鎖反応に使用されるオリゴヌクレオチドが配列番号：32、33、34、35、36、37、38、39または40からなる群より選択されるヌクレオチド配列を有する請求項24記載の方法。

【請求項26】 配列番号：32、33、34、35、36、37、38、39または40に示すヌクレオチド配列を有する単離ポリヌクレオチド。

【請求項27】 (a) 配列番号：32および34；

(b) 配列番号：33および34；

(c) 配列番号：35および37；

(d) 配列番号：36および37；

(e) 配列番号：38および40；または

(f) 配列番号：39および40；

に示すヌクレオチド配列を有するオリゴヌクレオチドを有してなるオリゴヌクレオチド対。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は単離されたヒトセリンプロテアーゼ（PSP1）ポリヌクレオチド、それらの相同体およびイソ体および多型性変種およびそれらの検出；本質的に純粋なPSP1タンパク質；ならびにPSP1ポリヌクレオチドおよびタンパク質の組成物および製造方法および使用に関する。

【0002】

【従来の技術および発明が解決しようとする課題】 プレセニリンの突然変異（PS-1およびPS-2）は家族性アルツハイマー病（FAD）の早期発病の全ケースの～95%（各々75%および20%）を占める。R. Sherringtonら、Nature 375: 754-760 (1995)；E. I. Rogaevら、Nature 376: 765-778 (1995)；およびE. Levy-Lahadら、Science 269: 973-977 (1995)を参照のこと。プレセニリンは高度な相同性を有する（同一性67%）マルチ膜スパンニングタンパク質（multi-membrane spanning protein）であるが、その機能は知られていない。

【0003】 全長46 kDaのPS-1タンパク質は、通常28 kDaおよび18 kDaの断片を有することが示されており；PS-2は同様に切断することが報告されている。M. Merckenら、FEBS Letters 389: 297-303 (1996)を参照のこと。この大きさの断片を生じる（複

数の) 切断部位は、エクソン8およびエクソン9によりコードされるタンパク質の領域にあると予想される。エクソン8はFADを招く突然変異のホットスポットである。従って、PSP-1のこの領域、およびプロセリナーゼプロテアーゼによるこの領域のPSP-1切断は、タンパク質の機能化に重要な事象である。PSP-1スパンニングエクソン8-11の領域は、プロテアーゼ、PSP1を特異的に結合し、PSP-1の内因性基質に対する活性および/またはPSP-1への結合能力はAD、前頭葉痴呆症、皮質性レーヴィ小体、パーキンソン病の痴呆症、急性および慢性の発作後の退化または頭部損傷、運動ニューロン疾患に見られるニューロン性退化、AIDS痴呆症ならびに慢性てんかんに関与する神経退化の病理学において重要であることが、本発明において示された。従って、PSP1に対応するヌクレオチドおよびアミノ酸配列の供給、PSP-1に結合するPSP1のモジュレーター、および/またはPSP1タンパク質溶解活性のモジュレーター、かかるモジュレーターの同定方法、ならびにかかる方法に有用な物質が必要とされる。

#### 【0004】

【課題を解決するための手段】従って、本発明の一つの態様は生物学的に活性なPSP1ポリペプチドをコードする単離ポリヌクレオチドである。本発明の別の態様は、以下の群から選択する単離ポリヌクレオチドである：

(a) 配列番号：24に示すヌクレオチド603から1979のヌクレオチド配列を有するPSP1-1をコードするポリヌクレオチド；および

(b) 実質的に配列番号：24に類似するポリヌクレオチド。

本発明の別の態様は、以下の群から選択する単離ポリヌクレオチドである：

(a) 配列番号：23に示すヌクレオチド603から1979のヌクレオチド配列を有するPSP1-2をコードするポリヌクレオチド；および

(b) 実質的に配列番号：23に類似するポリヌクレオチド。

本発明の別の態様は、以下の群から選択する単離ポリヌクレオチドである：

(a) 配列番号：26に示すヌクレオチド603から1979のヌクレオチド配列を有するPSP1-3をコードするポリヌクレオチド；および

(b) 実質的に配列番号：26に類似するポリヌクレオチド。

本発明の別の態様は、以下の群から選択する単離ポリヌクレオチドである：

(a) 配列番号：28に示すヌクレオチド603から1979のヌクレオチド配列を有するPSP1-4をコードするポリヌクレオチド；および

(b) 実質的に配列番号：28に類似するポリヌクレオ

チド。

【0005】さらなる態様において、本発明は上記で定義するポリヌクレオチドにおいて、ヌクレオチド672および1435がCおよびTから別個に選択される、任意の単離ポリヌクレオチドを提供し、本明細書において「多型性変種」と称する。本発明の別の態様は、本発明のポリヌクレオチドによりコードされる機能性ポリペプチドである。本発明の別の態様は、本発明またはD87258のポリヌクレオチドに結合する能力のある配列を有してなるアンチセンスオリゴヌクレオチドである。本発明の別の態様は、本発明またはD87258のポリヌクレオチドのモジュレーターである。

【0006】本発明の別の態様は、PSP1またはD87258の細胞性結合パートナーに対する結合に影響を及ぼすことにより、PSP1またはD87258の活性を調節する基質の存在に関して培地を評価する方法であって、以下の工程からなる：

(a) PSP1-1、PSP1-2、PSP1-3もしくはPSP1-4、またはD87258のアミノ酸配列を有するPSP1もしくはD87258タンパク質、またはそれらの機能的誘導体もしくは多型性変種および細胞性結合パートナーまたはそれらの合成類似体を提供すること；

(b) PSP1またはD87258タンパク質/細胞性結合パートナー複合体を形成できる条件下で、PSP1またはD87258活性を調節すると推測される試験物質とインキュベートすること；

(c) 複合体、遊離のPSPもしくはD87258タンパク質または遊離の細胞性結合パートナーの存在を評価すること；および

(d) 対照と比較して、該物質の効果を決定すること。

【0007】本発明の別の態様は、細胞性基質に及ぼすタンパク質溶解活性を阻害することにより、PSP1またはD87258の活性を調節する基質の存在に関して培地を評価する方法であって、以下の工程からなる：

(a) PSP1-1、PSP1-2、PSP1-3もしくはPSP1-4、またはD87258のアミノ酸配列を有するPSP1もしくはD87258タンパク質、またはそれらの機能的誘導体もしくは多型性変種および細胞性基質またはそれらの合成類似体を提供すること；

(b) PSP1酵素/基質複合体を形成し、続いて基質を切断できる条件下で、PSP1またはD87258活性を阻害すると推測される試験物質とインキュベートすること；

(c) タンパク質溶解により切断される基質の存在を評価すること；および

(d) 対照と比較して、該物質の効果を決定すること。

【0008】本発明の別の態様は、PSP1またはD87258タンパク質に直接的に結合することにより、PSP1またはD87258の活性を調節する物質の存在

を評価する方法であって、以下の工程からなる：

(a) PSP1-1、PSP1-2、PSP1-3もしくはPSP1-4、またはD87258のアミノ酸配列を有する標識化PSP1もしくはD87258タンパク質、またはそれらの機能的誘導体もしくは多型性変種を提供すること；

(b) モジュレーター候補物質と結合する固体支持体を提供すること；

(c) PSP1タンパク質／モジュレーター候補物質複合体を形成できる条件下で、標識化PSP1またはD87258タンパク質の混合物を支持体に結合したモジュレーター候補物質とインキュベートすること；

(d) 遊離の可溶性標識化PSP1またはD87258タンパク質から固体支持体を分離すること；

(e) 固体支持体に結合する標識化タンパク質の存在を評価すること；

(f) 標識化PSP1またはD87258タンパク質と複合した固体支持体を単離すること；および

(g) モジュレーター候補物質を同定すること。

【0009】本発明の別の態様は、本発明の方法により同定されるPSP1またはD87258タンパク質調節化合物である。本発明の別の態様は、本発明のPSP1またはD87258活性調節化合物の治療的有効量を、PSP1またはD87258活性を調節する必要性を有する患者に投与することからなる、かかる患者の治療方法である。

【0010】本発明の別の態様は、以下からなるPSP1またはD87258タンパク質欠乏に関連する状態の診断方法である：

(a) 個体からポリヌクレオチドサンプルを単離すること；

(b) PSP1またはD87258をコードするポリヌクレオチドサンプルおよび本発明のポリヌクレオチドを評価すること；および

(c) 任意の差異がPSP1またはD87258配列における突然変異を示す、ポリヌクレオチドサンプルおよびPSP1またはD87258ポリヌクレオチド間の差異を比較すること。

【0011】本発明の別の態様は、以下からなるPSP1またはD87258タンパク質機能不全に関連する状態の処置方法である：

(a) PSP1またはD87258タンパク質機能が欠損した患者から細胞を単離すること；

(b) PSP1またはD87258を発現する細胞に本発明またはD87258のポリヌクレオチドをトランスフェクトすることにより、細胞を変化させること；および

(c) 症状を緩和するために患者に細胞を再導入すること。

【0012】本発明の別の態様は、本発明のポリヌクレ

オチドをPSP1タンパク質機能欠損の患者に投与することからなる、PSP1またはD87258タンパク質機能不全に関与する状態の処置方法であって、PSP1またはD87258タンパク質を発現し、症状を緩和する方法である。本発明の別の態様は、PSP1もしくはD87258と免疫反応性のある抗体またはそれらの免疫原である。本発明の別の態様は、任意の細胞において本発明のポリヌクレオチドの発現能力のあるトランスジェニック非ヒト動物である。本発明の別の態様は、患者のサンプル中のPSP1またはD87258多型体を検出することからなる、患者における神経退化の遺伝的素因の測定方法である。本発明のさらなる別の態様は、配列番号：32、33、34、35、36、37、38、39、または40に示すヌクレオチド配列を有する単離ポリヌクレオチドである。

#### 【0013】

【発明の実施の形態】本明細書で用いる「PSP1ポリヌクレオチド」または「PSP1」なる用語は、PSP1および代替的にスプライス変種、すなわち相同体およびイソ体、ならびに多型性変種をコードするヌクレオチド配列を有してなるDNA分子を意味する。PSP1はヒトPSP1タンパク質のアミノ酸269-413を包含する領域に結合し、保存されたセリンプロテアーゼモチーフ (Ohno, I. ら、ジーンバンク受託番号D87258 (1996) ) を含有し、Lipinska ら、Nucl. Acids Res. 16:10053-10066 (1988) に記載されるイー・コリ (E. coli) セリンプロテアーゼ htrA および、本明細書でD87258と称する、IGF-結合モチーフを有する推定ヒトセリンプロテアーゼに相同性を呈する。PSP1-1配列は配列番号：24に示す。この配列のコード化領域は、配列番号：24のヌクレオチド603-1979を有してなる。PSP1-1コード化産物の458の誘導アミノ酸配列を配列番号：25に示す。

【0014】配列番号：30に示すPSP1-1配列は、ヌクレオチド672 (C/T) および1435 (C/T) で二つの多型性変種を含有し、24 (arg/cys) および278 (ala/val) の位置で代替アミノ酸基になり、両者共にヌクレオチド1-1540の保存領域にある。PSP1-1コード化産物の458の誘導アミノ酸配列を配列番号：31に示す。PSP1-2配列は配列番号：23に示す。この配列のコード化領域は、配列番号：23のヌクレオチド603-1979を有してなる。PSP1-2コード化産物の458の誘導アミノ酸配列を配列番号：8に示す。PSP1-3配列は配列番号：26に示す。この配列のコード化領域は、配列番号：26のヌクレオチド603-1736を有してなる。PSP1-3コード化産物の377の誘導アミノ酸配列を配列番号：27に示す。PSP1-4配列は配列番号：28に示す。この配列のコード化領域は、配列番号：28のヌクレオチド603-1913を

有してなる。PSP1-4コード化産物の436の誘導アミノ酸配列を配列番号：29に示す。

【0015】D87258配列は配列番号：17に示す。この配列のコード化領域は、配列番号：17のヌクレオチド49-1491を有してなる。D87258のコード化産物の480の誘導アミノ酸配列を配列番号：18に示す。配列番号：17に示したD87258配列は1325 (G/T) ヌクレオチドで多型性変種を含有し、213 (gly/val) の位置で代替えアミノ酸基になる。ジーンバンク受託番号D87258 (1996) の配列は単に1325Gと記載する。1325Tを有する新規ポリヌクレオチド多型体は、本明細書においてD87258 (1325T) と称し、213でバリンを有する新規にコードされた産物はD87258 (1325T) タンパク質である。新規ポリヌクレオチドD87258 (1325T) およびそれにコードされるタンパク質は、本明細書に記載される任意の組成物、使用および方法で、ならびにこれもまた本発明の一部を成す任意のかかる新規のポリペプチド、コードされるタンパク質、組成物、使用および方法において、PSP-1を置き換えることができる。

【0016】本明細書で用いる「機能性断片」なる用語は、具体的な遺伝子または遺伝子産物を修飾するために使用する場合、遺伝子の全長またはそれに関連する遺伝子産物に関わる、実質的に全ての生物学的機能を保持する遺伝子の全長以下の長さの部分またはかかる遺伝子産物を意味する。PSP1の機能的断片の例は、最少触媒領域である。特定の遺伝子または遺伝子産物の断片が機能的断片であるかどうかを決定するために、周知のヌクレオチド溶解法もしくはタンパク質溶解法により、またはポリメラーゼ連鎖反応により断片を生じ、その断片を記載の生物学的機能について試験する。本明細書で用いる「抗原」なる用語は、宿主の免疫系を刺激して、体液性および/または細胞性抗原特異性反応を引き起こす、1またはそれ以上のエピトープを含有する分子を意味する。この用語はまた本明細書で「免疫原」と互換的に用いられる。

【0017】本明細書で用いる「エピトープ」なる用語は、抗原または特異的抗体分子が結合するハプテン上の部位を意味する。この用語はまた本明細書で「抗原決定基」または「抗原決定部位」と互換的に用いられる。本明細書で用いる「モノクローナル抗体」は、一つの種

(例えばネズミ、ウサギ、ヤギ、ラット、ヒト等) に由来する抗体、および二つ (またはそれ以上) の種 (例えばキメラおよびヒト化抗体) に由来する抗体を包含すると理解される。RNAポリメラーゼが単一のmRNAに二つのコード化配列を転写し、次いでこのmRNAが両方のコード化配列に由来するアミノ酸を有する単一のポリペプチドに翻訳する場合、本明細書で用いるコード化配列は別のコード化配列に「機能的に連結する」。発現

した配列が最終的にプロセッシングされて望ましいタンパク質を産生する限り、コード化配列は互いに連続している必要はない。

【0018】本明細書で用いる「組換え」ポリペプチドは、組換えDNA技術により産生した、すなわち望ましいポリペプチドをコードする外因性DNA構築物により形質転換された細胞から産生したポリペプチドを意味する。「合成」ポリペプチドは化学的合成により製造したポリペプチドである。本明細書で用いる「レプリコン」は、インビボでDNA複製の自立的ユニットとして機能する；すなわち自身の制御下で複製できる任意の遺伝的要素 (例えばプラスミド、染色体、ウイルス) である。本明細書で用いる「ベクター」は、別のDNAセグメントを付着し、付着したセグメントが複製される様にしたプラスミド、ファージ、またはコスミド等のレプリコンである。本明細書で用いる「対照体」遺伝子は、本発明の野生型PSP1配列を意味し、遺伝子配列中にヌクレオチド置換体が存在するが、遺伝子産物の本質的な機能には影響しない、種々の配列多型体を包含すると考えられる。

【0019】本明細書で用いる「突然変異」遺伝子は、ヌクレオチド置換および/または削除および/または挿入が、遺伝子産物の本質的な機能を混乱させる、対照遺伝子とは異なるPSP1配列を意味する。本明細書で用いる特定のタンパク質のDNA「コード化配列」、または特定のタンパク質を「コードするヌクレオチド配列」は、適当な調節配列の制御下にある場合、ポリペプチドに転写および翻訳されるDNA配列である。本明細書で用いる「プロモーター配列」は、細胞内でRNAポリメラーゼに結合し、下流 (3' 方向) のコード化配列の転写を開始する能力のあるDNA調節領域である。本発明を定義する目的で、プロモーター配列は、コード化配列の翻訳開始コドン (例えばATG) により3' 末端で結合し、最小限の塩基またはバックグラウンドを超える検出可能なレベルで転写を開始するのに必要な要素を包含するように上流 (5' 方向) に伸長する。プロモーター配列内には、転写開始部位 (ヌクレアーゼS1のマッピングで便宜的に定義される) およびRNAポリメラーゼの結合に寄与するタンパク質結合領域 (コンセンサス配列) がある。真核生物プロモーターは、いつもではないが、しばしば、「TATA」ボックスおよび「CAV」ボックスを含有する。原核生物プロモーターは-10および-35コンセンサス配列に加えてシャインダーガルノ (Shine-Dalgarno) 配列を含有する。

【0020】本明細書で用いるDNA「制御配列」は総括的にプロモーター配列、リボソーム結合部位、ポリアデニル化信号、転写終止配列、上流制御領域、エンハンサー等を意味し、宿主細胞におけるコード化配列の発現 (すなわち転写および翻訳) を包括的に提供する。本明細書で用いるように、RNAポリメラーゼがプロモータ



一配列に結合し、コード化配列をmRNAに転写し、かかるmRNAをコード化配列によりコードされるポリペプチドに翻訳する場合、制御配列は細胞においてコード化配列の「発現を指示する」。本明細書で用いる「宿主細胞」は、外因性DNA配列により形質転換もしくはトランスフェクトされた、または形質転換もしくはトランスフェクトされ得る細胞である。

【0021】本明細書で用いるように、かかる外因性DNAが細胞膜内に導入された場合、細胞は外因性DNAにより「形質転換されている」。外因性DNAは、細胞のゲノムを構成する染色体DNAに組み込まれて（共有結合して）もしなくてもよい。原核生物および酵母では、例えば、外因性DNAはプラスミド等のエピソード要素上に保持され得る。真核細胞に関しては、安定して形質転換またはトランスフェクトされた細胞は、外因性DNAが染色体内に組み込まれ、娘細胞により染色体複製を介して遺伝される細胞である。この安定性は真核細胞がセルラインまたは外因性DNAを含有する娘細胞の集団からなるクローンを確立する能力により示す。

【0022】本明細書で用いる、「トランスフェクション」または「トランスフェクトされた」とは、細胞が外来性のDNAを取り込み、外来性のDNAを染色体に組み込む過程を意味する。トランスフェクションは例えば細胞がDNAを取り込む種々の技術（例えばリン酸カルシウム沈殿、電気泳動、リボソームの同化等）またはウイルスを用いてDNAを細胞に移すインフェクションにより達成できる。本明細書で用いる「標的細胞」は、別の細胞型（またはセルライン）で選択的にトランスフェクトされる細胞である。本明細書で用いる「クローン」は、単一の細胞または有糸分裂による共通の先祖に由来する細胞の集団である。「セルライン」は、多くの世代のためのインビトロでの安定した成長が可能な1次細胞のクローンである。

【0023】本明細書で用いるDNA構築物の「異型性」領域は、天然で他の分子に関連しては見出されない、別のDNA分子内またはその分子に付着した同定可能なDNAセグメントである。従って、異型性領域が遺伝子をコードする場合、遺伝子は通常源動物のゲノムで遺伝子をフランキングしないDNAによりフランキングされる。異型性コード化配列の別の例は、コード化配列自身が天然には見出されない構築物である（例えば天然の遺伝子とは異なるコドンを含む合成配列）。対立遺伝子変種または自然発生突然変異は本明細書で用いるDNAの異型性領域を生じない。本明細書で用いるポリペプチドの「モデュレーター」は、ポリペプチド機能に影響する物質、例えば酵素活性の阻害物質である。

【0024】本発明の一つの態様は、単離されたPSP1タンパク質をコードするポリヌクレオチドおよび実質的に類似の配列である。単離ポリヌクレオチド配列が適度に厳密な条件下で、配列番号：23、24、26もし

くは28にハイブリダイズできるか、またはこれらが配列番号：23、24、26もしくは28に同義的であるか、もしくは適度に厳密な条件下で、配列番号：23、24、26もしくは28にハイブリダイズできる配列に同義的であるDNA配列をコードする場合、単離ヌクレオチド配列は実質的に類似している。

【0025】「適度に厳密な条件」は当業者に理解される用語であり、例えばSambrookら、Molecular Cloning: A Laboratory Manual第2版、1巻101-104頁、コールド・スプリング・ハーバー・ラボラトリー・プレス（1989）に記載されている。適度に厳密な条件を用いる実例的なハイブリダイゼーション・プロトコールは以下のとおりである。ニトロセルロース・フィルターを6X SSPE、5Xデンハーツ溶液（Denhardt's solution）（溶液1あたりフィコール（Ficoll）10g、BSA 10gおよびポリビニルピロリドン10g）、0.005% SDSおよび100μg/ml tRNAを含有する溶液中65℃でプレハイブリダイズする。ハイブリダイゼーション・プローブは標識化、好ましくは放射性標識化（例えばBiosTAG-IT（登録商標）キット）する。次いでハイブリダイゼーションを65℃で約18時間実施する。フィルターを次いで2X SSCおよび0.5% SDSの溶液で室温で15分間で2回洗浄する。続いてフィルターを58度で洗浄し、空気乾燥し、増強スクリーンで-70℃で一晩X線フィルムに曝露する。

【0026】同義性DNA配列は、配列番号：8、25、27もしくは29、または適度に厳密な条件下で配列番号：8、25、27もしくは29にハイブリダイズできる配列によりコードされるタンパク質と同一のアミノ酸をコードするが、遺伝的コードの同義性のために、ヌクレオチドコード化配列に変種を有する。例えば、同義コドンUUCおよびUUUは共にアミノ酸フェニルアラニンをコードし、一方GGX（X=U、C、AまたはG）の四つのコドンは全てグリシンをコードする。また別に、実質的に類似する配列は、約70%、好ましくは約80%、最も好ましくは約90%PSP1とヌクレオチドの同一性を共有する、PSP1活性を有するタンパク質をコードするヌクレオチド配列であると定義する、すなわち全ヌクレオチド配列の少なくとも約70%が配列番号：23、24、26または28に適合する場合、PSP1活性を有するタンパク質をコードする配列は、任意の配列番号：23、24、26または28に実質的に類似する。実質的に類似するヌクレオチド配列は、ハイブリダイゼーションにより、または配列比較により同定できる。

【0027】本発明の単離ポリヌクレオチドの具体例には、DNA、ゲノムDNAおよびRNAがあり、好ましくはヒト由来のものである。PSP1タンパク質をコードする核酸分子を単離する方法は、当業界周知の方法を

用いて、天然のまたは人工的に設計されたプローブでゲノムまたはcDNAライブラリーをプローブする。例えば「Current Protocols in Molecular Biology」Ausubelら(編)、グリーン・パブリッシング・アソシエーション・アンド・ジョーン・ウィレイ・インターサイエンス、ニューヨーク、1989、1992を参照のこと。通常の当業者は、少なくとも15の近接するヌクレオチドを有してなる配列番号：23、24、26もしくは28またはそれらの断片が、特に有用なプローブであることを理解している。かかるプローブを、プローブの同定を容易にするために、分析的に検出可能な物質で標識化でき、標識化することが好ましいことも理解している。有用な物質には放射性アイソトープ、蛍光染料または検出可能な産物の形成を触媒できる酵素等があるが、これらに限定するものではない。プローブにより当業者がヒト、哺乳動物またはその他の動物に由来するPSP1タンパク質をコードするゲノムDNA、cDNAまたはRNAポリヌクレオチドの相補的コピーを単離し、またはかかる由来源に関連配列例えばファミリーのさらなるメンバー、タイプおよび/またはサブタイプ、例えば転写調節および制御要素ならびにその他の安定性、プロセッシング、翻訳および本明細書に開示するコード化配列に相關する5'および/または3'領域からの組織特異性決定領域に関してスクリーニングすることが可能になるが、全て必要以上に実験する必要はない。

【0028】本発明の別の態様は、本発明のポリヌクレオチドによりコードされる機能性ポリペプチドおよび実質的に類似のポリペプチドである。本発明の機能性ポリペプチドの一つの具体例は、配列番号：8、25、27または29に示すアミノ酸配列を有するPSP1タンパク質である。実質的に類似のポリペプチド配列は、PSP1活性を有し、約50%、好ましくは約70%、最も好ましくは約90%PSP1と同等なアミノ酸を共有するポリペプチド配列、即ちPSP1活性を有するポリペプチドを示す配列が実質的に任意の配列番号：8、24、26または28に類似し、配列の全アミノ酸の少なくとも約50%が配列番号：8、25、27または29に適合する場合である。実質的に類似するポリペプチド配列は、タンパク質溶解性消化、ゲル電気泳動、マイクロシーケンシングおよび/または配列比較、(例えばウィスコンシン大学ジェネティクス・コンピューター・グループから利用可能なGAPアルゴリズムの使用による)等により同定できる。

【0029】本発明の別の態様は、本質的に純粋なPSP1タンパク質を製造する方法である。さらに別の態様は、本発明の製造方法により産生されたPSP1タンパク質である。このタンパク質は配列番号：8、25、27または29に示すアミノ酸配列を有し、同一の機能を有する実質的に類似したアミノ酸配列を有する変種を包含する。本発明のタンパク質は、タンパク質の発現およ

びそれらの回収を促進する条件下で、本発明のポリヌクレオチドをコードするベクターを含有する組換え宿主細胞を培養する、組換え遺伝子操作技術により製造するのが好ましい。

【0030】単離ポリヌクレオチド、とりわけDNAは、DNAを遺伝子発現に必要な、必須発現制御領域、例えば調節領域に機能的に連結することにより、発現ベクターに導入できる。ベクターは適当な宿主細胞、例えば細菌等の原核生物、または酵母もしくは哺乳動物細胞等の真核生物に、当業界周知の方法で導入できる。Ausubelら、(上記に引用)を参照のこと。製造し単離した望ましいタンパク質のコード化配列を、任意の適当なベクターまたはレプリコンにクローン化できる。多くのクローニングベクターが当業者に周知であり、適当なクローニングベクターの選別は、選択の問題である。クローニング用の組換えDNAベクターおよび形質転換できる宿主細胞の例としては、バクテリオファージ(イー・コリ：E.coli)、pBR322(イー・コリ：E.coli)、pACYC177(イー・コリ：E.coli)、pK1230(グラム陰性細菌)、pGV1106(グラム陰性細菌)、pLAFR1(グラム陰性細菌)、pME290(イー・コリ：E.coli以外のグラム陰性細菌)、pHV14(イー・コリ：E.coliおよびバチルス：サブチリス：Bacillus subtilis)、pBD9(バチルス：Bacillus)、pIJ61(ストレプトミセス：Streptomyces)、pUC6(ストレプトミセス：Streptomyces)、YIP5(サッカロミセス：Saccharomyces)、バキュロウイルス(baculovirus)昆虫細胞系、ドロソフィラ(Drosophila)昆虫系、YCP19(サッカロミセス：Saccharomyces)ならびにpSV2ネオ(哺乳動物細胞)等があるが、これらに限定するものではない。一般的には「DNAクローニング」：I&I巻、Gloverら編、IRLプレス・オックスフォード(1985)(1987)；およびT.Maniatisら、「Molecular Cloning」コールド・スプリング・ハーバー・ラボラトリー(1982)を参照のこと。

【0031】遺伝子は制御要素、例えばプロモーター、リボソーム結合部位(細菌性発現用)および所望によりオペレーター等の制御下に置き、望ましいタンパク質をコードするDNA配列を、発現構築物を含有するベクターにより形質転換された宿主細胞中のRNAに転写するようにする。コード化配列は単一のペプチドまたはリーダー配列を含有してもしなくともよい。本発明のタンパク質は、例えばイー・コリ(E.coli)tacプロモーターまたはA遺伝子(spa)プロモーターおよび信号配列を用いて発現できる。リーダー配列は翻訳後プロセッシング中の細菌宿主により除去され得る。例えば米国特許第4431739号；第4425437号および第4338397号を参照のこと。

【0032】制御配列に加え、宿主細胞の成長に相対し

てタンパク質配列の発現を調節できる調節配列を加えるのが望ましい。調節配列は当業者に周知である。遺伝子の発現を化学的または物理的刺激に反応してスイッチを入れたり切ったりさせるものを例示すると、調節化合物の存在、または種々の温度もしくは代謝条件等がある。他の型の調節要素は、ベクター中にも存在でき、エンハンサー配列等がある。

【0033】特定のコード化配列を適当な調節配列を有するベクター中に置き、コード化配列の位置および配向は、制御配列に関して、コード化配列が制御配列の制御下で転写される、すなわち制御配列でDNA分子に結合するRNAポリメラーゼがコード化配列を転写するようにして、発現ベクターを構築する。目的の特定の抗原をコードする配列の修飾は、この目標を達成するために望ましい。例えば、適当な配向で制御配列に結合する、すなわち解読枠を維持するために配列を修飾することが必要になりうる場合もある。制御配列および他の調節配列はベクター、例えば上記のクローニングベクター等に挿入する前にコード化配列にライゲートできる。また別に、コード化配列はすでに制御配列および適当な制限部位を含有している発現ベクターに直接的にクローン化できる。

【0034】PSP1タンパク質の突然変種または類似体を産生するのが望ましい場合もある。突然変種または類似体は、タンパク質をコードする配列の一部の削除、配列の挿入および/または配列内の1またはそれ以上のヌクレオチドの置換により、製造できる。ヌクレオチド配列を修飾する技術、例えば部位指向性突然変異誘発は、当業者に周知である。例えばT.Maniatisら、上記で引用；「DNA Cloning」I & II巻；および「Nucleic Acid Hybridization」上記で引用を参照のこと。発現系および宿主選別に応じて、本発明のタンパク質は、目的のタンパク質を発現するための条件下で、上記の発現ベクターにより形質転換した成長宿主細胞により産生される。好ましい哺乳動物細胞には、ヒト腎胚細胞(293)、サル腎細胞、繊維芽(COS)細胞、チャイニーズ・ハムスター卵巣(CHO)細胞、ドロソフィラ(Drosophila)またはネズミL-細胞等がある。発現系がタンパク質を成長培地中に分泌する場合、タンパク質は培地から直接的に精製できる。タンパク質が分泌しない場合、細胞ライゼートから単離するかまたは細胞膜分画から回収する。適当な成長条件および回収方法の選別は、当業者の範疇である。

【0035】本発明のタンパク質を同定するための別の方法は、イー・コリ(E.coli)を形質転換するために得られたクローンをを用いて遺伝子ライブラリーを構築し、PSP1に対するポリクローナル血清またはモノクローナル抗体を用いて個々のコロニーをブルーおよびスクリーニングすることによる。本発明のタンパク質はまた、化学的合成法、例えば自動ペプチド合成器での固相ペプ

チド合成により、既知アミノ酸配列または目的の遺伝子のDNA配列に由来するアミノ酸配列を用いて産生できる。このような配列は当業者に周知である。本発明のタンパク質、または少なくとも一つのエピトープを有してなるそれらの免疫原断片は、本明細書に開示するアミノ酸配列に対応するエピトープに指向するポリクローナルおよびモノクローナルの両方の抗体を産生するために使用できる。ポリクローナル抗体が望ましい場合、選別した哺乳動物、例えばマウス、ウサギ、ヤギまたはウマは、本発明のタンパク質、またはその断片または突然変異タンパク質で免疫化する。免疫化した動物からの血清を回収し、既知の方法に準じて処理する。血清ポリクローナル抗体は、イムノアフィニティー・クロマトグラフィーまたはその他の既知の方法により精製できる。

【0036】本発明のタンパク質およびそれらの免疫原断片に対するモノクローナル抗体もまた、当業者により容易に産生できる。ハイブリドーマ技術を用いる一般的なモノクローナル抗体製造のための方法がよく知られている。普遍的な抗体産生セルラインは、細胞融合により、また腫瘍形成DNAでのBリンパ球の直接的な形質転換またはエプステイン・バー( Epstein-Barr) ウイルスでのトランスフェクション等のその他の技術によっても作ることができる。例えばM.Schreierら、「Hybridoma Techniques」(1980)；Hammerlingら、「Monoclonal Antibodies and T-cell Hybridomas」(1981)；Kennettら、「Monoclonal Antibodies」(1980)；および米国特許第4341761号；第4399121号；4427783号；第444887号；第4452570号；第4466917号；第4472500号；第4491632号および第4493890号を参照のこと。目的の抗原に対して産生したモノクローナル抗体のパネル、またはその断片は種々の特性、すなわちイソ体、エピトープ、親和性等をスクリーニングできる。モノクローナル抗体は、免疫親和性技術を用いた、それらを指向する個々の抗原の精製に有用である。また別に、目的のモノクローナル抗体をコードする遺伝子は当業者に周知のPCR技術によりハイブリドーマより単離でき、適当なベクターにクローン化および発現できる。本発明の抗体は、ポリクローナルでもモノクローナルでも、イムノアッセイ、RIA、ELISA等において試薬として用いることができるという点でさらなる利用性がある。本発明の抗体は、分析的に検出可能な試薬、例えば放射性アイソトープ、蛍光分子または酵素で標識化できる。

【0037】ヒト以外の動物の可変領域がヒト定常領域に結合または融合しているキメラ抗体(例えばLiuら、Proc.Natl.Acad.Sci.USA、84:3439(1987)を参照のこと。)は、分析において、または治療的にも有用である。治療用モノクローナル抗体はJohnsら、Nature 321:522(1986)；Verhoeyenら、Science 239:1534(198

8) ; Kabatら、J. Immunol. 147:1709 (1991) ; Queenら、Proc. Natl. Acad. Sci. USA、86:10029 (1989) ; Gormanら、Proc. Natl. Acad. Sci. USA、88:34181 (1991) ; およびHodgsonら、Bio/Technology 9:421 (1991) に記載されるように「ヒト化」されるのが好ましい。本発明の別の態様は、本発明のポリペプチドまたはD87258のモジュレーターである。物質によるPSP1またはD87258の機能的調節には、機能の部分的阻害から完全阻害、例えばタンパク質溶解活性の阻害、同等な機能および機能の増強等がある。本発明のモジュレーターの具体例には、ペプチド、オリゴヌクレオチドおよびペプチド擬似物質等の小さな有機物質等がある。本発明のモジュレーターは、AD等の神経退化の全ての型に治療的または予防的に有用である。PSP1またはD87258の、その他の外因性物質に相対するタンパク質溶解活性のモジュレーターは、別の型のヒトの病態の処置にも有用である。

【0038】本発明の別の態様は、本発明のポリペプチドに結合できる配列を有してなるアンチセンスオリゴヌクレオチドである。合成オリゴヌクレオチドまたは関連するアンチセンス化学的構造性類似体は、PSP1またはD87258タンパク質をコードする標的核酸を認識し、それに特異的に結合し、転写を阻止する様に設計できることは、当業者に理解され得る。一般的には、Cohen, J. S., Trends in Pharm. Sci. 10:435 (1989) およびWeintraub, H. M., Scientific American, 1990年1月、40頁を参照のこと。本発明の別の態様は、PSP1またはD87258タンパク質の細胞性結合パートナーへの結合に影響することにより、PSP1またはD87258タンパク質機能を調節する物質の存在に関して培地を評価するための方法である。モジュレーターの例としては、ペプチド、ペプチド擬似物質等の小さな有機分子等があるが、これらに限定するものではない。PSP1またはD87258タンパク質は、細胞性結合パートナーまたはそれらの合成類似体を伴うPSP1 (配列番号: 8、25、27または29) もしくはD87258タンパク質 (配列番号: 18) またはそれらの機能的断片のアミノ酸配列を備えている。混合物を、PSP1またはD87258遺伝子産物/細胞性結合パートナー複合体を形成できる条件下で、PSP1またはD87258活性を調節する可能性のある試験物質とインキュベートする。複合体、遊離のPSP1またはD87258タンパク質または遊離の細胞性結合パートナーの存在に関してアッセイを実施し、結果を対照と比較して試験物質の効果を決定する。

【0039】本発明の別の態様は、細胞性物質に及ぼすタンパク質溶解活性を阻害することにより、PSP1またはD87258タンパク質の機能を調節する物質の存在に関して培地を評価する方法である。モジュレーターの例としては、ペプチド、ペプチド擬似物質等の小さな

有機分子等があるが、これらに限定するものではない。細胞性物質にはPSP1、PSP2、APPまたはその他の物質等がある。PSP1またはD87258タンパク質は、細胞性基質またはそれらの合成類似体を伴うPSP1 (配列番号: 8、25、27または29) もしくはD87258タンパク質 (配列番号: 18) またはそれらの機能的断片のアミノ酸配列を備えている。混合物を、PSP1またはD87258酵素/基質複合体を形成し、続いて基質を切断できる条件下で、PSP1またはD87258活性を阻害すると推測される試験物質とインキュベートする。

【0040】本発明の別の態様は、PSP1またはD87258タンパク質に直接的に結合することにより、PSP1またはD87258の活性を調節する物質の存在を評価する方法である。モジュレーターの例としてはペプチド、ペプチド擬似物質等の小さな有機分子等があるが、これらに限定するものではない。モジュレーター候補物質はLamら、Nature 354:82 (1991) またはBurbaumら、Proc. Natl. Acad. Sci. USA 92:6027 (1995) に開示される技術等により、固体支持体上で合成し、固体支持体に結合したモジュレーター候補物質を提供する。標識化PSP1またはD87258タンパク質は、PSP1 (配列番号: 8、25、27または29) もしくはD87258タンパク質 (配列番号: 18) またはそれらの機能的誘導体のアミノ酸配列を備えている。標識を例示すると、直接的に結合する蛍光または着色染料、ビオチン、放射性アイソトープまたはエピトープタグ等があり、適切な抗体により検出可能である。固体支持体に結合したモジュレーター候補物質、および標識化PSP1またはD87258タンパク質の混合物は、PSP1またはD87258タンパク質/モジュレーター候補物質複合体を形成できる条件下でインキュベートする。固体支持体は遊離の可溶性標識化PSP1またはD87258タンパク質から分離する。固体支持体に結合した標識化タンパク質の存在に関してアッセイを実施する。標識化タンパク質と複合体形成する固体支持体を単離し、モジュレーター候補物質は、当業者に周知の技術、例えばTOF-SIMS法 (Brummelら、Science 264:399-402 (1994)) により同定する。

【0041】PSP1またはD87258機能の調節はプレセニリン切断、その他のタンパク質の切断またはβA4産生に影響を及ぼすことが予期される。このように同定された任意のモジュレーターは、FADおよびAD等の神経退化の処置および予防に治療用として有用であると予期される。さらに、PSP1またはD87258は、相互作用するタンパク質を単離するために用いることができ、この相互作用は妨害の標的となり得る。PSP1またはD87258およびその他の因子との間のタンパク質-タンパク質相互作用の阻害物質により、PSP1またはD87258活性を調節するための医薬物的

質が開発された。タンパク質-タンパク質相互作用、例えばPSP1またはD87258遺伝子産物/結合パートナー複合体の相互作用当を評価する方法、およびPSP1またはD87258と相互作用するタンパク質を単離する方法は、当業者に周知である。以下に記載する方法を用いると、通常の当業者は必要以上の実験を必要とせずにこれらの目的を達成することが可能となる。

【0042】酵母2ハイブリッド系 (Yeast two-hybrid system) は、インピグで転写活性化物質の活性を再構成して、最初の試験タンパク質および2番目の試験タンパク質間の相互作用を検出する方法を提供する。この方法は米国特許第5283173号に開示されており、試薬はクロンテックおよびストラタジーンから入手可能である。簡単に説明すると、PSP1 cDNAをGal4またはLexA転写因子DNA結合領域に融合し、酵母細胞中に発現する。目的の細胞より入手したcDNAライブラリー・メンバーは、Gal4のトランスアクチベーション (transactivation) 領域または別のトランスアクチベーション領域に融合する。PSP1と相互作用できるタンパク質を発現するcDNAクローンにより、転写因子活性、例えばGal4およびリポーター遺伝子発現例えばGal1-lacZのトランスアクチベーションが再構成される。

【0043】別の方法は、λgt11、λZAP (ストラタジーン) または組換えPSP1と等価なcDNA発現ライブラリーをスクリーニングする方法である。組換えPSP1タンパク質またはそれらの断片は、小さなペプチドタグ、例えばFLAG、HSVまたはGSTに融合する。ペプチドタグは、心筋クレアチンキナーゼ等のキナーゼのための便宜的なリン酸化部位を有することができるか、またはビオチニル化できる。組換えPSP1は<sup>32</sup>Pでリン酸化できるか、または非標識化したものを用いてストレプトアビジンもしくはタグに対する抗体で検出できる。λgt11 cDNA発現ライブラリーは目的の細胞から作り、組換えPSP1とインキュベートし、洗浄し、PSP1と相互作用するcDNAクローンを単離する。T. Maniatisら、(上記で引用) を参照のこと。

【0044】別の方法は、cDNAがベクターに、哺乳動物プロモーターおよびポリペプチドアデニル化部位の間でクローン化した、哺乳動物発現ライブラリーをスクリーニングし、COS細胞または293細胞に一時的にトランスフェクトし、次いで48時間後に標識化、好ましくはヨウ素化PSP1を有する細胞を同定し、洗浄し、インキュベートして、結合タンパク質を検出し、オートラジオグラフィで結合性PSP1を検出する方法である (Simsら、Science 241:585-589 (1988) およびMcMahanら、EMBO J. 10:2821-2832 (1991) を参照のこと)。この方法で、目的の結合性タンパク質をコードするcDNAを含有するcDNAプールを選別でき、目的

のcDNAは、各プールをさらにサブ分割することにより単離でき、次いで一時的なトランスフェクション、結合およびオートラジオグラフィのサイクルを続ける。また別に、目的のcDNAは全cDNAライブラリーを哺乳動物にトランスフェクトし、プレートに結合したPSP1を含有する皿上で細胞をパニングすることにより単離できる。洗浄後に付着する細胞を溶解し、プラスミドDNAを単離し、細菌中で増幅し、トランスフェクションおよびパニングのサイクルを単一のクローンが得られるまで繰り返す (Seedら、Proc. Natl. Acad. Sci. USA 84:3365 (1987) およびAruffoら、EMBO J. 6:3313 (1987) を参照のこと)。結合タンパク質が分泌される場合、一度結合または中和アッセイは一時的なトランスフェクト細胞からの上澄を評価するために確立されると、かかるcDNAは類似のプーリング法により入手できる。上澄の一般的なスクリーニング方法は、Wongら、Science 228:810-815 (1985) に開示されている。

【0045】また別の方法は、PSP1と相互作用するタンパク質を直接的に細胞から単離する方法である。PSP1とGSTまたは小さなペプチドタグとの融合タンパク質を作り、ビーズ上に固定化する。生合成的に標識化した、または非標識化タンパク質は目的の細胞からの抽出物を調製し、ビーズと共にインキュベートし、緩衝液で洗浄する。PSP1と相互作用するタンパク質は、ビーズから特異的に溶出し、SDS-PAGEにより分析する。結合パートナー主要アミノ酸配列データは、マイクロシーケンシングにより入手する。所望により、細胞を細胞性タンパク質の機能的反応、例えばチロシンリン酸化等を誘起する物質と処理する。このような物質の例としては、成長因子またはサイトカイン例えばエリスロポイエチンまたはインターロイキン-3等がある。また別の方法は、免疫親和性精製法である。組換えPSP1は標識化または非標識化細胞抽出物と共にインキュベートし、抗PSP1抗体で免疫沈殿させる。免疫沈殿物はプロテインA-セファロースで回収し、SDS-PAGEで分析する。非標識化タンパク質は、ビオチニル化により標識化し、ストレプトアビジン含有するSDSゲル上で検出する。結合パートナータンパク質は、マイクロシーケンシングにより分析する。さらに、標準的な生化学的精製工程は、当業者に周知であり、マイクロシーケンシングの前に用いることができる。

【0046】さらに別の方法は、結合パートナーのためのペプチドライブラリーのスクリーニング法である。組換えタグ化または標識化PSP1は、PSP1と相互作用するペプチドまたはリン酸ペプチドライブラリーからペプチドを選別するために用いる。ペプチドのマイクロシーケンシングにより相互作用タンパク質に見出され得るコンセンサスペプチド配列が同定できるようになる。通常の当業者に周知のこれらのまたはその他の方法の任意の方法により同定したPSP1またはD87258結合パー

トナー、および上記のこれらの推定結合パートナーは、本発明のアッセイに用いることができる。PSP1またはD87258/結合パートナー複合体の存在の評価は、例えば酵母2-ハイブリッド系 (yeast two-hybrid system)、ELISAまたは複合体に特異的な抗体を用いたイムノアッセイにより達成できる。PSP1またはD87258/結合パートナー相互作用の形成を妨害または阻害する試験物質が存在する場合、複合体の量は、試験物質を欠いた対照と比較して減少する。

【0047】遊離PSP1もしくはD87258または結合パートナーに関するアッセイは、例えばELISAもしくは特異的抗体を用いるイムノアッセイにより、または放射性標識化PSP1もしくはD87258を細胞もしくは細胞膜と共にインキュベートし、続いて遠心もしくはろ過分離工程により達成される。PSP1またはD87258/結合パートナー相互作用の形成を妨害または阻害する試験物質が存在する場合、遊離のPSP1もしくはD87258または遊離の結合パートナーの量が、試験物質を欠く対照と比較して増加する。

【0048】本発明の別の態様は、本発明のPSP1またはD87258モデュレーターの有効量、および医薬的に許容できる担体を有してなる医薬組成物である。本発明のモデュレーターの医薬組成物は、非経口投与用すなわち皮下、筋肉内、もしくは静脈内投与用または経口投与用に製造できる。非経口用組成物は、通常本発明のモデュレーターの溶液、または許容できる担体、好ましくは水性担体に溶解したそれらのカクテルから成る。種々の水性担体には、例えば水、緩衝液、0.4%生理食塩水、0.3%グリシン等が用いられる。これらの溶液は無菌で、一般的には微粒子は含まれない。これらの溶液は通常周知の無菌化技術により無菌化できる。組成物は医薬的に許容できる、生理学的状態に近づけるために必要な補助物質、例えばpH調節物質および緩衝化物質等を含有できる。かかる医薬的製剤における本発明のモデュレーターの濃度は、非常に広範に、すなわち重量で約0.5%以下から、通常または少なくとも約1%で、15または20%ほどまでに変化でき、選択した特定の投与様式に準じて、主に液体容量、粘度等に基づいて選択する。

【0049】このように、本発明のモデュレーターの筋肉内注射用の医薬的組成物は、1mlの無菌緩衝液、および本発明のタンパク質50mgを含有するように製造できる。同様に、本発明のモデュレーターの静脈内還流用の医薬的組成物は、250mlの無菌リンガー溶液 (Ringer's solution)、および本発明のタンパク質モデュレーター150mgを含有するように製造できる。非経口投与可能な組成物の実際的な製造方法は周知であるか、または当業者に明白であり、さらに詳細には、例えばRemington's Pharmaceutical Science、第15編、マック・パブリッシング・カンパニー、イーストン、ベ

ンシルバニアに記載されている。

【0050】医者は最も適した本発明の治療用物質の投与量を決定し、これは投与形態および選択する特定の化合物により変更され、さらに処置中の特定の患者に応じて変更される。一般的に、医者は化合物の至適投与量よりも実質的に少ない投与量で処置を開始し、少量ずつ投与量を増加し、このような状況下で至適効果が得られるまで増加させることを望む。組成物が経口的に投与される場合、非経口投与で少量で得られたのと同じの効果を出すために、活性物質の量はより多量必要とされる。治療用投与量は一般的に0.1~1000mg/日であり、高用量であるが、幾つかの異なる投与量ユニットで投与することができる。

【0051】患者の状態に応じて、本発明の医薬的組成物は、予防的および/または治療的処置のために投与できる。治療的適用の場合、本発明の化合物を含有する組成物またはそれらのカクテルを、既に疾病に冒された患者に、疾病または合併症を治療または少なくとも部分的に抑止するのに十分な量投与する。予防的に投与する場合、本発明の化合物を含有する組成物またはそれらのカクテルを、未だ疾病状態にはない患者に、患者の疾病に対する抵抗性を増強するために投与する。医薬的組成物の単回または多数回投与は、処置する医師が選択する投与量レベルおよびパターンで実施できる。いかなる場合も、本発明の医薬的組成物は、本発明のモデュレーターが、患者を効果的に処置するのに十分な量を提供すべきである。

【0052】さらに、欠陥遺伝子の遺伝により引き起こされる疾病もある。これらの遺伝子は欠陥遺伝子の配列を正常の遺伝子配列と比較することにより検出できる。PSP1またはD87258遺伝子の突然変異を担う個体は、種々の技術によりDNAレベルで検出できる。診断用の核酸 (ゲノムDNA、mRNA等) は、患者の細胞、例えば血液、尿、唾液、または組織生検例えば絨毛膜絨毛サンプリングまたは羊水細胞の除去、および剖検素材等から入手できる。ゲノムDNAは直接検出に用いることができ、またはPCR、リガーゼ連鎖反応 (LCR)、鎖置換増幅 (SDA) 等を用いて分析前に酵素的に増幅できる。例えばSaikiら、Nature324:163-166 (1986)、Bejら、Crit. Rev. Biochem. Molec. Biol. 26:301-334 (1991)、Birkenmeyerら、J. Virol. Meth. 35:117-126 (1991)、Van Brunt、J. Bio/Technology 8:291-294 (1990)を参照のこと。RNAまたはcDNAはまた同じ目的のために用いることができる。例を挙げると、本発明の核酸に相補的なPCRプライマーは、PSP1またはD87258突然変異体を同定および分析するために使用できる。例えば、削除および挿入は、正常のPSP1またはD87258ゲノタイプと比較して、増幅産物の大きさの変化により検出できる。点突然変異は、増幅DNAを本発明の放射性標識化PSP1またはD872

58 RNAに、また別に本発明の放射性標識化PSP1またはD87258アンチセンスDNA配列にハイブリダイズさせることにより同定できる。完全に対合する配列は、RNAアーゼA消化または融解温度(Tm)の差により誤対合二重らせんより区別できる。このような診断法は特に、出生前試験および新生児試験にも有用である。

【0053】さらに、点突然変異ならびに対照遺伝子および「突然変異」遺伝子間の他の配列の差異は、他の周知の技術、例えば直接的なDNAシーケンシング、一本鎖立体配位多型性により同定できる。Oritaら、Genomics 5:874-879 (1989)を参照のこと。例えば、シーケンシングプライマーは、二本鎖PCR産物、または改変PCRにより生じた一本鎖鋳型分子と共に用いる。放射性標識化ヌクレオチドを用いる通常の方法により、または蛍光タグを用いる自動シーケンシング法により、配列決定を行う。クローン化したDNAセグメントは、特異的DNAセグメントを検出するためのプローブとしても使用することができる。この方法の感度は、PCRと組み合わせることにより非常に増強される。さらに、点突然変異および他の配列変種、例えば多型体は上記の方法、例えば単一のヌクレオチドにより異なる配列のPCR増幅のための対立遺伝子特異性オリゴヌクレオチドの使用により、検出できる。配列番号: 32、33、34、35、36、37、38、39および40に示す配列を有するオリゴヌクレオチドがこのような方法において有用である。これらの方法は、患者のサンプル中のPSP1またはD87258の多型性を検出することにより、患者における神経退化的遺伝的素因の決定するのに有用である。多型性はPSP1の672ヌクレオチドで、PSP1の1435ヌクレオチドで、またはD87258の1325ヌクレオチドで検出されるのが好ましい。多型性はPCRで検出するのが好ましく; 多型性は配列番号: 32、33、34、35、36、37、38、39および40から成る群から選択したヌクレオチド配列を有するオリゴヌクレオチドでのPCRにより検出するのが最も好ましい。神経退化的素因はアルツハイマー病に対して決定されるのが好ましい。

【0054】DNA配列の差異に基づく遺伝的試験は、変成物質を伴うか、または伴わないゲル中のDNA断片の電気泳動の可動性の変化により検出できる。小さな配列の欠損および削除は、高分解能ゲル電気泳動により可視化できる。配列が異なるDNA断片は、変成ホルムアミド・グラジエントゲル上で区別でき、ゲル中の異なるDNA断片の可動性は、特異的融解または部分融解温度に準じて、ゲル中の異なる位置で遅延する。例えばMyersら、Science 230:1242 (1985)を参照のこと。さらに、配列の変化、とりわけ小さな欠損は、非変成ゲル電気泳動例えばヘテロ二重らせん電気泳動での、DNAヘテロ二重らせんの移動パターンの変化により検出でき

る。例えばNagamineら、Am. J. Hum. Genet. 45:337-339 (1989)を参照のこと。特異的な位置での配列変化は、ヌクレアーゼ保護アッセイ、例えばRNAアーゼおよびS1保護、またはCottonら、Proc. Natl. Acad. Sci. USA 85:4397-4401 (1985)に開示される化学的切断法によっても示すことができる。このように、特異的DNA配列の検出は、例えばハイブリダイゼーション(例えばヘテロ二重らせん電気泳動、Whiteら、Genomics 12:301-306 (1992)を参照のこと)、RNAアーゼ保護(例えばMyersら、Science 230:1242 (1985))、化学的切断(例えばCottonら、Proc. Natl. Acad. Sci. USA 85:4397-4401 (1985))、直接的DNAシーケンシング、または制限酵素の使用(例えば、制限断片の数および大きさのバリエーションが、エンドヌクレアーゼ制限配列を創り出すまたは破壊する挿入、欠失、ヌクレオチドの反復の存在および任意のその他の突然変異を示し得る、制限断片長多型性(restriction fragment length polymorphisms: RFLP))により達成できる。ゲノムDNAのサザンブロッティングを用いて大きな(すなわち100塩基対より大きい)欠損および挿入をも同定できる。

【0055】通常のゲル電気泳動およびDNAシーケンシングに加えて、突然変異例えば微小欠失、異数性、転座、逆位もまたインサイトー分析により検出できる。例えばKellerら、DNA Probes、第2版、ストックトン・プレス、ニューヨーク、ニューヨーク州、米国(1993)を参照のこと。すなわち、細胞におけるDNAまたはRNA配列は、単離および/または膜への固定化をすることなく突然変異を分析できる。蛍光インサイトーハイブリダイゼーション(FISH)は、現在最も一般的に適用される方法であり、FISHに関して、非常に多くの文献がある。例えばTrachuckら、Science 250:559-562 (1990)、およびTraskら、Trends, Genet. 7:149-154 (1991)を参照のこと。従って、PSP1またはD87258遺伝子の構造に基づく核酸を用いて、遺伝的突然変異の診断用試験を開発することができる。

【0056】さらに、mRNAにおける変化により検出できる遺伝子発現の変化によりもたらされる、またはこの変化により特徴づけられる疾患もある。また別に、PSP1またはD87258遺伝子を対照として用いて、PSP1またはD87258 mRNAの発現レベルが増加または低下する個体を、例えばノーザンブロッティングまたはインサイトーハイブリダイゼーションにより同定できる。適当なハイブリダイゼーション条件の規定は当業者の範疇である。例えば「Current Protocols in Mol. Biol.」I & II巻、ウィレイ・インターサイエンス、Ausbelら、(編)(1992)を参照のこと。プロービング技術は当業界で周知であり、プローブの大きさは広範に変化できるが、プローブは少なくとも15ヌクレオチドであるのが好ましいということは理解されよう。このようなプローブはプローブの同定を容易にするため



に、分析的に検出可能な試薬で標識化するのが好ましいということも理解されよう。有用な試薬には放射性アイソトープ、蛍光染料または検出可能な産物の形成を触媒できる酵素等があるが、これらに限定するものではない。一般的な法則としては、ハイブリダイゼーション条件をより厳密な状態にするほど、回収される遺伝子はより正確に相関するようになる。

【0057】プレセニリン生化学におけるPSP1またはD87258の推定される役割は遺伝子治療であり、本発明の別の態様を為す。「遺伝子治療」は遺伝子補充を意味し、目的の遺伝子の対照体コピーをさらに患者の細胞に挿入する。その結果、対照遺伝子によりコードされたタンパク質が欠陥を修復し、細胞の機能を正常にし、従って疾病の症状を緩和する。対照体コピーはPSP1またはD87258遺伝子の野生型形態、または内因性PSP1またはD87258の活性を調節するタンパク質またはペプチドをコードする遺伝子である。本発明の遺伝子治療は、インビボまたはエクソビボで可能である。エクソビボ遺伝子治療では、患者の細胞の単離および精製、治療用遺伝子の導入、および遺伝的に変化した細胞の患者への再導入が必要とされる。複製不能ウイルス例えば修飾レトロウイルスを用いて、かかる細胞に治療用PSP1またはD87258遺伝子を導入できる。例えば、マウスモロニー白血病ウイルス(MMLV)は臨床の遺伝子治療試行において周知のベクターである。例えばBoris-Lauerieら、Curr. Opin. Genet. Dev. 3:102-109 (1993)を参照のこと。

【0058】これに比して、インビボ遺伝子治療は患者の細胞の単離および精製を必要としない。治療用遺伝子は、典型的なものでは、患者に投与するために例えばリポソームに、またはアデノウイルス等の複製不能ウイルス(Berkner, K. L., Curr. Top. Microbiol. Immunol. 158: 39-66 (1992)に記載)もしくはアデノ関連ウイルス(AAV)ベクター(Muzyczka, N., Curr. Top. Microbiol. Immunol. 158:97-129 (1992)および米国特許第5252479号に記載)に「パッケージング」する。別の方法は治療用の遺伝子を直接的に血流または筋肉組織に注射する、「裸のDNA (naked DNA)」の投与である。別の方法は、DNAで被覆した金粒子を用いた微粒子ボンバードメントにより、標的組織に治療用遺伝子を

導入する、「裸のDNA」の投与である。

【0059】本発明の遺伝子治療に有用な細胞型には、リンパ球、肝細胞、筋芽細胞、繊維芽細胞、眼の任意の細胞例えば網膜細胞、上皮および内皮細胞等がある。細胞は処置すべき患者から採取したTリンパ球、肝細胞、眼の任意の細胞または呼吸器もしくは肺上皮細胞であるのが好ましい。肺上皮細胞のトランスフェクションは、リポソーム中のDNAベクターの調製物、DNAタンパク質複合体または複製不能アデノウイルスを霧状にして吸入して行うことができる。例えば米国特許第5240

846号を参照のこと。

【0060】本発明の別の態様は、任意の細胞に本発明のポリヌクレオチドまたはD87258を発現する能力のある、トランスジェニック、非ヒト哺乳動物である。トランスジェニック、非ヒト動物は、適当な受精卵または宿主の胚を本発明のポリヌクレオチドで、D87258でまたはヒト疾患に見出される突然変種でトランスフェクトすることにより入手できる。例えば米国特許第4736866号;第5175385号;第5175384号および第5175386号を参照のこと。得られたトランスジェニック動物はPSP1またはD87258遺伝子機能の研究のためのモデルとして用いることができる。とりわけ有用なトランスジェニック動物は、PSP1またはD87258タンパク質の発現に関連する検出可能な表現型を呈示するものである。医薬品開発の候補物質は次いで関連する表現型を逆転するまたは悪化させる能力をスクリーニングする。本発明は具体的で非限定的な実施例を参考に記載する。

【0061】

【実施例】

実施例1ー PS-1結合パートナーPSP1の同定  
PS-1アミノ酸配列の269-413の残基(配列番号:10)をコードするPS-1 cDNA(ジーンバンク受託番号L42110)(配列番号:9)の一部を、オリゴヌクレオチドプライマー5'-CGGAATTC CGTATGCTGGTTGAAACA-3'(配列番号:11)および5'-CGGGATCCTCAGGC TACGAAACAGGCTAT-3'(配列番号:12)でPCR増幅した。産生物をEcoRIおよびBamHIで消化し、pEG202(Golemisら、Current Protocols in Molecular Biology、ジョーン・ウィレイ & サンズ、ニューヨーク(1994))にクローン化した。得られたプラスミド、pCC352は、DNA結合タンパク質、LexAをPS-1のアミノ酸269-413にフレーム内融合した融合タンパク質をコードする。親ベクター、pEG202はアルコール脱水素酵素(ADHI)プロモーターを用いて、LexA融合タンパク質および選別マーカーとしてHIS3を発現する酵母発現ベクターである。自動DNAシーケンサー(アプライド・バイオシステムズ、インコーポレーティッド)を用いる配列分析により、増幅領域が正しい配列を有しており、LexAにフレーム内融合したことを確認した。

【0062】2ハイブリッドスクリーニングに用いた全体的方法、プラスミドおよび株はGolemisら、(上記に引用)に詳細に記載されている。酵母株EGY48(MATa, trp1, his3, ura3, 6ops-L EU2)は、プラスミドpCC352およびpSH18-34で同時形質転換した。形質転換体は、ウラシルおよびヒスチジンを欠く完全最少培地を用いて選別した。プラスミドpSH18-34は、8個のLexAオペレ



ーター部位が、LacZ遺伝子および選別マーカースとしてURA3の発現を誘導する最少GAL1プロモーターの上流に位置する、酵母発現ベクターである。LexA-PS-1融合体の全長の合成は、LexAに対して指向するポリクローナル抗血清を用いて酵母抽出物をウェスタンブロット分析により確認した。LexA-PS-1融合体単独ではLEU2もLacZリポーター株をも活性化できなかった。さらに、LexA-PS-1融合体の核への侵入能力およびDNA結合能力はレプレッションアッセイを用いて確認した。

【0063】LexA-PS-1融合体およびpSH18-34 (CCY321) を含有する株は、ライブラリー・スケール形質転換プロトコルを用いて、プラスミドpJG4-5中ヒト胎児脳cDNAライブラリー (クロンテック) で形質転換した。このライブラリープラスミドは、TRP1選別マーカースを含有し、SV40核局在化配列、酸ブロッブ (acid blob) B42、およびヘマグルチニンエピトープタグを含有するカセットに、融合体 (AD融合体) としてcDNAを発現させる。Gyurisら、Cell 75:791-803 (1993) を参照のこと。この融合体の発現は、ガラクトース誘導プロモーターGAL1の制御下にある。形質転換反応物はウラシル、ヒスチジンおよびトリプトファンを欠く完全最少培地にプレートした。約 $4.5 \times 10^6$ の個々の形質転換体が得られ、プールし、凍結させた。各々の主要コロニーを選別中に再度プレートされたことを確認するために、 $2 \times 10^7$ 生存細胞 (個々の形質転換体の約3倍の数) をウラシル、ヒスチジン、トリプトファンおよびロイシンを欠き、AD融合の発現を誘起するための炭素源としてガラクトース/ラフィノースを含む最少培地にプレートした。30℃で3および4日間成長させ、生じたコロニーをウラシル、ヒスチジンおよびトリプトファンを欠く完全最少培地に採った。潜在的相互作用融合タンパク質を含有するコロニーは次にガラクトース依存性およびLacZ発現を試験した。ガラクトース依存の形でLEU2およびLacZリポーターの両方を活性化するこれらの単離体は、陽性と考えられ、さらに試験を実施した。プラスミドを酵母から単離し、イー・コリ (E. coli) KC8株を形質転換するのに用い、トリプトファンを欠く最少イー・コリ (E. coli) 培地上で成長させることによりAD融合プラスミドを選別した。潜在的相互作用融合体を含有する各々のAD融合プラスミドは、CCY321を形質転換するために使用した。幾つかの形質転換体はガラクトース依存性LEU2およびLacZ活性化のスクリーニングに供した。相互作用が特異的であることを確認するために、各AD融合プラスミドの22の非相関性LexA融合タンパク質と相互作用する能力を試験した。スクリーニングの第二ラウンドを通過し、LexA-PS-1融合体と特異的に相互作用するAD融合プラスミドを同定した。

#### 【0064】実施例2- PSP1 cDNAクローニングおよび配列分析

AD融合プラスミドを制限消化分析および上記のシーケンシングに供した。相互作用融合タンパク質cDNAの一つの配列分析により、配列番号: 1の2の位置でGGAで始まり、523の位置でTGAで終わる173アミノ酸 (配列番号: 2) タンパク質をコードする519ヌクレオチドオープンリーディングフレーム (配列番号: 1) が示された。cDNA配列またはイー・コリ

10 (E. coli) セリンプロテアーゼhtrAの一部に相同性を示す誘導アミノ酸配列を有するBLASTXおよびBLASTNアルゴリズムを用いたジーンバンク・サーチは、上記で引用するLipinskaら、に記載されている (配列番号: 13および14)。この新規cDNAはPSP1と称する。cDNAのより大きな部分を得るために、配列番号: 1の83-106塩基対を表すオリゴヌクレオチド5'-CTGGATGGGGAGGTGATTGGGAGTG-3' (配列番号: 15) を用いて、ジレントラッパーcDNA陽性選別システム (ギブコBRL) を使用して、スーパースクリプト (Superscript) ヒト脳cDNAライブラリー (ギブコBRL) をスクリーニングした。Innisら、PCR Protocols: A Guide to Methods and Applications、アカデミック・プレス、サンディエゴ、カリフォルニア州 (1990) およびSambrookら、Molecular Cloning: A Laboratory Manual、第2版、コールド・スプリング・ハーバー・プレス、コールド・スプリング・ハーバー、ニューヨーク州 (1989) に記載されるように、全形細胞PCRまたは標準的なハイブリダイゼーション条件を用いてコロニーをスクリーニングした。PSP1を含有するこれらの単離体は制限消化分析およびシーケンシングに供した。最も長いクローンである配列番号: 3および配列番号: 5は全体をシーケンシングした。

【0065】配列番号: 3の配列分析により、配列番号: 3の1の位置でCCCで始まり、972の位置でTGAで終わる323アミノ酸 (配列番号: 4) タンパク質をコードする969ヌクレオチドオープンリーディングフレームが示された。配列番号: 5の配列分析により、配列番号: 5の1の位置でCTTで始まり、1272の位置でTGAで終わる423アミノ酸 (配列番号: 6) タンパク質をコードする1500ヌクレオチドオープンリーディングフレームが示された。スクリーニングの第二ラウンドは、配列番号: 5の5-28塩基対を表す、オリゴヌクレオチド5'-GTCTCTGGGCCCCGGTTGTCTGTTG-3' (配列番号: 16) を用いて実施した; ライブラリーおよびスクリーニングプロトコルは変更しない。スクリーニングの第二ラウンドでは、単離した消化した配列番号: 7は最長のcDNAクローンを含有する。配列番号: 7の配列分析により、配列番号: 7の251の位置でATGで始ま

り、1627の位置でTGAで終わる458アミノ酸  
(配列番号: 8) タンパク質をコードする1374ヌク  
レオチドオープンリーディングフレームが示された。し  
かしながら、配列番号: 7は潜在的な開始コドンから上  
流には停止コドンを有さない。予想される出発コドンが  
本物であることを確認するために、「マラソン・レディ  
ー (Marathon Ready)」ヒト脳cDNA (クロンテッ  
ク) およびプライマーのネスト・セット (nested set)  
を用いて、5'ヌクレオチド配列を5' RACEで伸長  
させた。配列番号: 7特異的プライマー5'-CCAA 10  
CAGACAACCGGGCCAGAGACT-3'  
(配列番号: 20) および5'アンカープライマー-1  
(クロンテック) を、最初のPCR増幅に使用し、配列  
番号: 7特異的プライマー5'-TGCCCTCCTCG  
CCCGCCCTACTCAGA-3' (配列番号: 2  
1) および5'アンカープライマー-2 (クロンテッ  
ク) を、2番目のPCR増幅に使用した。PCR産物  
はpCR2.1 (インビトロジェン) にT/Aクローン  
化した。不安定な (staggered) 5'末端を有する18  
の単離体を分析し、587ヌクレオチドの5'コンセン  
サス配列を生じた (配列番号: 22)。コンセンサス配  
列 (配列番号: 23) を生じるために、配列番号: 22  
および配列番号: 7を整列化すると、ヌクレオチドの2  
25の位置にフレーム内停止コドンがあり、最初のメチ  
オニンが配列番号: 7の予想される位置に対応すること  
が示された。この遺伝子をPSP1-2と称する。

【0066】遺伝子のコンセンサス配列の全長はPSP  
1-1 (配列番号: 24および25)、PSP1-3  
(配列番号: 26および27) およびPSP1-4 (配  
列番号: 28および29) と称し、5'コンセンサス配  
列 (配列番号: 22)、他の部分的OSP1クローン、  
ならびに各々配列番号: 7、3および5を伴う整列化に  
より生じた。PSP1-1 (配列番号: 25) の誘導ア  
ミノ酸のイー・コリ (*E. coli*) htrA (配列番号:  
14) への整列化は、BESTFITアルゴリズム (ウ  
ィスコンシン大学ジェネティクス・コンピューター・  
グループ) を用いて達成した。アミノ酸レベルで約55%  
の類似性および33.5%の同一性が観察され、図1 (上  
部、PSP1-1; 下部、イー・コリ (*E. coli*) htrA)  
に示す。全てのセリンタンパク質溶解酵素に保存  
される、臨界ヒスチジンおよびセリンモチーフGXSG  
は、PSP1-1の各々198および304-308  
のアミノ酸位置に存在し、太字で示す。アミノ酸数は整  
列化した配列の左および右に示す。

【0067】ギャップ・クリエーション (gap creatio  
n) およびエクステンション・ペナルティー (extension  
penalties) が各々5.0および0.3である、PILEUP  
およびPRETTYアルゴリズム (ウィスコンシン大学  
ジェネティクス・コンピューター・グループ) を用い  
たPSP1-2、PSP1-1、PSP1-3およびP 50

SP1-4のヌクレオチド配列比較は図2-10に示  
す。整列化の結果、整列のヌクレオチド1541の位置  
で、PSP1-2およびPSP1-1は225塩基対を  
欠失し、PSP1-4は195塩基対を欠失することが  
示された。同一の整列内の1942のヌクレオチド位置  
で、PSP1-4はPSP1-2、PSP1-1および  
PSP1-3に存在する96塩基対を欠いている。各欠  
失部位の接合部には、スプライス部位コンセンサス配列  
AGGまたはTGG (太字で示す) があり、これらの交  
換形態 (alternate forms) は代替的なスプライシン  
グに因るものであることを示唆している。Mount, S., Nu  
cl. Acids. Res. 10:458-472 (1982) を参照のこと。15  
41の位置での見かけのスプライシングにより、PSP  
1-3に存在する停止コドン (図2-10で下線を付  
す) が除去される。さらに、PSP1-2およびPSP  
1-1は整列の672の位置で単一のヌクレオチドが相  
違ふ。PSP1-2は、システインをコードするTG  
Cコドンを作るTをこの位置で含有し、一方PSP1-  
1はシステインをコードするCGCコドンを作るCをこ  
の位置で含有する。PS1-1 (配列番号: 24) の上  
記で引用したOhnoらの推定ヒトセリンプロテアーゼ (配  
列番号: 17) に対するヌクレオチド配列比較により、  
GAPアルゴリズムを用いると49%の同一性が、BE  
STFITアルゴリズムを用いると65%の同一性が示  
された (データは示していない)。PSP1-1 (配列  
番号: 25) の誘導アミノ酸配列の、上記で引用したOh  
noらのD87258プロテアーゼ (配列番号: 18) に  
対する整列化は、BESTFITアルゴリズムを用いて  
達成し、図11 (上部、PSP1-1; 下部、Ohnoらの  
D87258プロテアーゼ) に示す。アミノ酸レベルで  
約46%の同一性が観察された。

【0068】実施例3- PSP1の組織分布  
ノーザン分析を実施して、ヒト組織内のPSP1mRN  
Aの分布を測定した。PSP1配列に対して指向する3  
0塩基オリゴヌクレオチドプローブを使用した (5'-  
ATGCTGAACATCGGGAAAGCTTGGT  
TCTCG-3') (配列番号: 19)。複数のヒト組  
織 (クロンテック #7750-1、#7760-1、  
および#7755-1) からのmRNAを含有するノー  
ザンブロットを、厳密な条件下でこのプローブとハイブ  
リダイズした。約1.9キロ塩基の主要なバンドは、試験  
した: 心臓、脳、肺、胎盤、肝臓、骨格筋、腎臓、脾  
臓、扁桃、尾状核、脳梁、海馬、黒質、視床下核、刺  
傷、小脳、大脳皮質、髄質、脊髄、後頭極、前頭葉、側  
頭極および被殻の全ての領域で検出された。PSP1m  
RNAはまたアルツハイマー病の脳でも検出された。

【0069】実施例4- PSP1多型性の検出  
PSP1オリゴヌクレオチド1AFC、1AFTおよび  
1ARは、ArgをCysにアミノ酸変換させるヌクレ  
オチド672 (シチジンからチミン) での多型性を検出

するために設計した。対立遺伝子特異性オリゴヌクレオチド (ASO) 1AFCおよび1AFTは、これらの3'末端塩基は別にして、同等であり、多型性をスクリーニングするための特異性を提供する。

1AFC: CAT CCG GCA TTG TTA  
GCT CTG C22メル (配列番号: 32)

1AFT: CAT CCG GCA TTG TTA  
GCT CTG T22メル (配列番号: 33)

1AR: CAA TAG CTG CAT CAG T  
TT GAA TG23メル (配列番号: 34)

オリゴヌクレオチド対 (1AFC+1AR、または1AFT+1AR) を以下の条件下でPCRに使用した: 94℃で40秒、60℃で30秒、1Uklentaq1 (ゲンパック・リミテッド)、50mM Tris-C1 pH9.1、16mM硫酸アンモニウム、3.5mM MgCl<sub>2</sub>、150μgml<sup>-1</sup>BSAおよび由来は未知のヒトゲノム25ngを含有する一回の反応で35サイクル。オリゴヌクレオチドの各対はゲノムDNAの12のランダムサンプルに対して試験し、産生物を4%寒天ゲル (ギブコー-BRL) で電気泳動した。95塩基対の予期される産生物は、これらの個体が、この多型生のために異型接合体であることを示す12個のDNAのうち8個における両方のASOで認められた。1AFCオリゴヌクレオチドのみで2個のDNAを増幅し、この位置でシスチジンを有する対立遺伝子に関して同型接合体である。1AFTオリゴヌクレオチドのみで2個のDNAを増幅し、この位置でチミンを有する対立遺伝子に関して同型接合体である。PSP1オリゴヌクレオチド1BFC、1BFTおよび1BRは、AlaをValにアミノ酸変化させるヌクレオチド1435 (シチジンからチミン) での多型性を検出するために設計した。

1BFC: TGG CGG GCT TTG GGG  
GGC ATT C22メル (配列番号: 35)

1BFT: TGG CGG GCT TTG GGG  
GGC ATT T22メル (配列番号: 36)

1BR: GAC GTC AGC AGG GCC C  
GG AGG TC23メル (配列番号: 37)

オリゴヌクレオチド対 (1BFC+1BR、または1BFT+1BR) を以下の条件下でPCRに使用した: 94℃で40秒、67℃で30秒、1Uklentaq1 (ゲンパック・リミテッド)、50mM Tris-C1 pH9.1、16mM硫酸アンモニウム、3.5mM M\*

配列番号: 1

GGGACTCCCC CAAACCAATG TGAATACAT TCAAAGTATG GCAGCTATTG ATTTTGGAAA	60
CTCTGGAGGT CCCCTGGTTA ACCTGGATGG GGAGGTGATT GGAGTGAACA CCATGAAGGT	120
CACAGCTGGA ATCTCCTTTG CCATCCCTTC TGATCGTCTT CGAGAGTTTC TGCATCGTGG	180
GGAAAAGAAG AATTCCTCCT CCGGAATCAG TGGGTCCAG CGGCGTACA TTGGGGTGAT	240
GATGCTGACC CTGAGTCCCA GCATCCTTGC TGAAGTACAG CTTCGAGAAC CAAGCTTTCC	300
CGATGTTTCA CATGGTGATC TCATCCATAA AGTCATCCTG GGCTCCCTG CACACCGGGC	360
TGGTCTGCGG CCTGGTGATG TGATTTTGGC CATTGGGGAG CAGATGGTAC AAAATGCTGA	420

\*gCl<sub>2</sub>、150μgml<sup>-1</sup>BSAおよび由来は未知のヒトゲノム25ngを含有する一回の反応で35サイクル。オリゴヌクレオチドの各対はゲノムDNAの12のランダムサンプルに対して試験し、産生物を4%寒天ゲル (ギブコー-BRL) で電気泳動した。75塩基対の予期される産生物は、その他の3個体が、この位置で異なる対立遺伝子を有することを示す12個のサンプルのうち9個で、1BFT ASOを用いて認められた。

【0070】実施例5- D87258多型性の検出  
オリゴヌクレオチド2AFC、2AFTおよび2ARは、GlyをValにアミノ酸変化させるヌクレオチド1325 (グアニンからチミン) での多型性を検出するために設計した。

2AFC: GAT ACC CCA GCA GAA  
GCT GG20メル (配列番号: 38)

2AFT: GAT ACC CCA GCA GAA  
GCT GT20メル (配列番号: 39)

2AR: GCT GAC ATC ATT GGC G  
GA GAC21メル (配列番号: 40)

オリゴヌクレオチド対 (2AFC+2AR、または2AFT+2AR) を以下の条件下でPCRに使用した: 94℃で40秒、62℃で30秒、1Uklentaq1 (ジンパック・リミテッド)、50mM Tris-C1 pH9.1、16mM硫酸アンモニウム、3.5mM MgCl<sub>2</sub>、150μgml<sup>-1</sup>BSAおよび由来は未知のヒトゲノム25ngを含有する一回の反応で35サイクル。オリゴヌクレオチドの各対はゲノムDNAの12のランダムサンプルに対して試験し、産生物を4%寒天ゲル (ギブコー-BRL) で電気泳動した。2AFT ASOは約1000塩基対の一つのバンドを生じた。予期されるバンドは90塩基対である。恐らく、より大きなバンドの存在はオリゴヌクレオチド2ARおよび2AFTによりフランキングされる領域のイントロンの存在に因るものである。2AFTで増幅したサンプルの全てにおいてバンドが観察され、このことはチミンを含有する対立遺伝子が12個体全てに存在することを示している。本発明は、本発明の精神または本質的な属性から離脱することなく、他の具体的な形態に具象化でき、従って、特許請求の範囲のみならず、明細書も、本発明の範囲を示す参考とすべきである。

【0071】

【配列表】

33 34  
 AGATGTTTAT GAAGCTGTTT GAACCCAATC CCAGTTGGCA GTGCAGATCC GGCGGGGACG 480  
 AGAAACACTG ACCTTATATG TGACCCCTGA GGTACACAGAA TGAATAGATC ACCAAGAGTA 540  
 TGAGGCTCCT GCTCTGATTT CCTCCTTGCC TTTCTGGCTG AGGTTCTGAG GGCACCGAGA 600  
 CAGAGGGTTA AATGAACCAG TGGGGGCAGG TCCCTCCAAC CACCAGCACT GACTCCTGGG 660  
 CTCTGAAGAA TCACAGAAAC ACTTTTATA TAAATAAAA TTATACCTAG CAACAAAAAA 720  
 AAAAAAAAAA AA 732

【0072】

配列番号：2

Gly Leu Pro Gln Thr Asn Val Glu Tyr Ile Gln Thr Asp Ala Ala Ile  
 1 5 10 15  
 Asp Phe Gly Asn Ser Gly Gly Pro Leu Val Asn Leu Asp Gly Glu Val  
 20 25 30  
 Ile Gly Val Asn Thr Met Lys Val Thr Ala Gly Ile Ser Phe Ala Ile  
 35 40 45  
 Pro Ser Asp Arg Leu Arg Glu Phe Leu His Arg Gly Glu Lys Lys Asn  
 50 55 60  
 Ser Ser Ser Gly Ile Ser Gly Ser Gln Arg Arg Tyr Ile Gly Val Met  
 65 70 75 80  
 Met Leu Thr Leu Ser Pro Ser Ile Leu Ala Glu Leu Gln Leu Arg Glu  
 85 90 95  
 Pro Ser Phe Pro Asp Val Gln His Gly Val Leu Ile His Lys Val Ile  
 100 105 110  
 Leu Gly Ser Pro Ala His Arg Ala Gly Leu Arg Pro Gly Asp Val Ile  
 115 120 125  
 Leu Ala Ile Gly Glu Gln Met Val Gln Asn Ala Glu Asp Val Tyr Glu  
 130 135 140  
 Ala Val Arg Thr Gln Ser Gln Leu Ala Val Gln Ile Arg Arg Gly Arg  
 145 150 155 160  
 Glu Thr Leu Thr Leu Tyr Val Thr Pro Glu Val Thr Glu  
 165 170

【0073】

配列番号：3

CCCAGTCTCT GGGCCCGGTT GTCTGTTGGG GTCAGTGAAC CCCGAGCATG CCTGACGTCT 60  
 GGGACCCCGG GTCCCGGGC ACAACTGACT GCGGTGACCC CAGATACCAG GACCCGGGAG 120  
 GCCTCAGAGA ACTCTGGAAC CCGTTCCGCG GCGTGGCTGG CCGTGGCGCT GGGCGCTGGG 180  
 GGGGCAGTGC TGTGTTGTTT GTGGGGCGGG GGTCCGGGTC CTCGGCCCGT CCTCGCCGCC 240  
 GTCCCTAGCC CGCCGCCCGC TTCTCCCGCG AGTCAGTACA ACTTCATCGC AGATGTGGTG 300  
 GAGAAGACAG CACCTGCCGT GGTCTATATC GAGATCCTGG ACCGGCACCC TTTCTTGGGC 360  
 CGCGAGGTCC CTATCTCGAA CGGCTCAGGA TTCGTGGTGG CTGCCGATGG GCTCATTGTC 420  
 ACCAACGCCC ATGTGGTGGC TGATCGGCGC ACAGTCCGTG TGAGACTGCT AAGCGGGCAG 480  
 ACGTATGAGG CCGTGGTCAC AGCTGTGGAT CCCGTGGCAG ACATCGCAAC GCTGAGGATT 540  
 CAGACTAAGG AGCCTCTCCC CACGCTGCCT CTGGGACGCT CAGCTGATGT CCGGCAAGGG 600  
 GAGTTTGTG TTGCCATGGG AAGTCCCTTT GCACTGCAGA ACACGATCAC ATCCGGCATT 660  
 GTTAGCTCTG CTCAGCGTCC AGCCAGAGAC CTGGGACTCC CCCAAACCAA TGTGGAATAC 720  
 ATTCAAACTG ATGCAGCTAT TGATTTTGA AACTCTGGAG GTCCCTGGT TAACCTGGTG 780  
 AGTGAGACAT CCTTCCTTCC AAGAATCCCT GCCCCAGGTC AGTGTGGGAA GGTAGGTTT 840  
 CCCCTAATTC AAGGATGTTT GGTCAAGTTT CTGAGCAGTT CTTTGTGGC TATCTCTCAA 900  
 TATCCAACCA GATCTCCCA AACTTGCTG GTACTTTTGT TCGGGTGGCC CCATCCCTTA 960  
 CTATTTGTTT AGGCTAGGGA ACTGGGGGCT GTATCCCTGC AGGATGGGGA GGTGATTGGA 1020  
 GTGAACACCA TGAAGTCAC AGCTGGAATC TCCTTTGCCA TCCCTTCTGA TCGTCTTCCA 1080

35

36

GAGTTTCTGC ATCGTGGGGA AAAGAAGAAT TCCTCCTCCG GAATCAGTGG GTCCCAGCGG 1140  
 CGCTACATTG GGGTGATGAT GCTGACCCTG AGTCCCAGCA TCCTTGCTGA ACTACAGCTT 1200  
 CGAGAACCAA GCTTTCCCGA TGTTTCAGCAT GGTGTACTCA TCCATAAAGT CATCCTGGGC 1260  
 TCCCCTGCAC ACCGGGCTGG TCTGCGGCCT GGTGATGTGA TTTTGGCCAT TGGGGAGCAG 1320  
 ATGGTACAAA ATGCTGAAGA TGTATTATGAA GCTGTTTCGAA CCCAATCCCA GTTGGCAGTG 1380  
 CAGATCCGGC GGGGACGAGA AACACTGACC TTATATGTGA CCCCTGAGGT CACAGAATGA 1440  
 ATAGATCACC AAGAGTATGA GGCTCCTGCT CTGATTTCCT CCTTGCCTTT CTGGCTGAGG 1500  
 TTCTGAGGGC ACCGAGACAG AGGGTTAAAT GAACCACTGG GGCAGGTCC CTCCAACCAC 1560  
 CAGCACTGAC TCCTGGGCTC TGAAGAATCA CAGAAACACT TTTTATATAA AATAAAATTA 1620  
 TACCTAGCAA CATATTATAG TAAAAAATGA GGTGGGAGGG CTGGATCTTT TCCCCACCA 1680  
 AAAGGCTAGA GGTAAAGCTG TATCCCCCTA AACTTAGGGG AGATACTGGA GCTGACCATC 1740  
 CTGACCTCCT ATTAAGAAA ATGAGCTGCT GAAAAA AAAA 1787

【0074】

配列番号：4

Pro Ser Leu Trp Ala Arg Leu Ser Val Gly Val Thr Glu Pro Arg Ala  
 1 5 10 15  
 Cys Leu Thr Ser Gly Thr Pro Gly Pro Arg Ala Gln Leu Thr Ala Val  
 20 25 30  
 Thr Pro Asp Thr Arg Thr Arg Glu Ala Ser Glu Asn Ser Gly Thr Arg  
 35 40 45  
 Ser Arg Ala Trp Leu Ala Val Ala Leu Gly Ala Gly Gly Ala Val Leu  
 50 55 60  
 Leu Leu Leu Trp Gly Gly Gly Arg Gly Pro Pro Ala Val Leu Ala Ala  
 65 70 75 80  
 Val Pro Ser Pro Pro Pro Ala Ser Pro Arg Ser Gln Tyr Asn Phe Ile  
 85 90 95  
 Ala Asp Val Val Glu Lys Thr Ala Pro Ala Val Val Tyr Ile Glu Ile  
 100 105 110  
 Leu Asp Arg His Pro Phe Leu Gly Arg Glu Val Pro Ile Ser Asn Gly  
 115 120 125  
 Ser Gly Phe Val Val Ala Ala Asp Gly Leu Ile Val Thr Asn Ala His  
 130 135 140  
 Val Val Ala Asp Arg Arg Arg Val Arg Val Arg Leu Leu Ser Gly Asp  
 145 150 155 160  
 Thr Tyr Glu Ala Val Val Thr Ala Val Asp Pro Val Ala Asp Ile Ala  
 165 170 175  
 Thr Leu Arg Ile Gln Thr Lys Glu Pro Leu Pro Thr Leu Pro Leu Gly  
 180 185 190  
 Arg Ser Ala Asp Val Arg Gln Gly Glu Phe Val Val Ala Met Gly Ser  
 195 200 205  
 Pro Phe Ala Leu Gln Asn Thr Ile Thr Ser Gly Ile Val Ser Ser Ala  
 210 215 220  
 Gln Arg Pro Ala Arg Asp Leu Gly Leu Pro Gln Thr Asn Val Glu Tyr  
 225 230 235 240  
 Ile Gln Thr Asp Ala Ala Ile Asp Phe Gly Asn Ser Gly Gly Pro Leu  
 245 250 255  
 Val Asn Leu Val Ser Glu Thr Ser Phe Leu Pro Arg Ile Pro Ala Pro  
 260 265 270  
 Gly Gln Cys Gly Lys Gly Arg Phe Pro Leu Ile Gln Gly Cys Leu Val  
 275 280 285

37 38  
 Lys Phe Leu Ser Ser Ser Leu Leu Ala Ile Ser Gln Tyr Pro Thr Arg  
 290 295 300  
 Ser Pro Gln His Leu Leu Val Leu Leu Phe Gly Cys Pro His Pro Leu  
 305 310 315 320  
 Leu Phe Val

【0075】

配列番号：5

CTTCGGGCAT GCGGGGCTTT GGGGGGCATT CGCTGGGGGA GGAGACCCCG TTTGACCCCT 60  
 GACCTCCGGG CCCTGCTGAC GTCAGGAACT TCTGACCCCG GGGCCCGAGT GACTTATGGG 120  
 ACCCCCAGTC TCTGGGCCCG GTTGTCTGTT GGGGTCACTG AACCCCGAGC ATGCCTGACG 180  
 TCTGGGACCC CGGTCCCGG GGCACAACCTG ACTGCGGTGA CCCCAGATAC CAGGACCCGG 240  
 GAGGCTCAG AGAACTCTGG AACCCGTTCC CGCGCGTGCC TGGCGGTGGC GCTGGGCGCT 300  
 GGGGGGGCAG TGCTGTTGTT GTTGTGGGGC GGGGCTCGGG GTCCTCCGGC CGTCCTCGCC 360  
 GCGCTCCCTA GCGCGCCGCC CGCTTCTCCC CGGAGTCAGT ACAACTTCAT CGCAGATGTG 420  
 GTGAGAAGA CAGCACCTGC CGTGGTCTAT ATCGAGATCC TGGACCGGCA CCCTTTCTTG 480  
 GGCCGGGAGG TCCCTATCTC GAACGGCTCA GGATTCGTGG TGGCTGCCGA TGGGCTCATT 540  
 GTCACCAACG CCCATGTGGT GGCTGATCGG CGCAGAGTCC GTGTGAGACT GCTAAGCGGC 600  
 GACACGTATG AGGCCGTGGT CACAGCTGTG GATCCCGTGG CAGACATCGC AACGCTGAGG 660  
 ATTCAGACTA AGGAGCCTCT CCCCACGCTG CCTCTGGGAC GCTCAGCTGA-TGTCCGGCAA 720  
 GGGGAGTTTG TTGTTGCCAT GGGAACTCCC TTTGCACTGC AGAACACGAT CACATCCGGC 780  
 ATTGTTAGCT CTGCTCAGCG TCCAGCCAGA GACCTGGGAC TCCCCCAAAC CAATGTGGAA 840  
 TACATTCAAA CTGATGCAGC TATTGATTTT GGAACTCTG GAGGTCCCCT GGTAAACCTG 900  
 GCTAGGGAAC TGGGGGCTGT ATCCCTGCAG GATGGGAGG TGATTGGAGT GAACACCATG 960  
 AAGGTCACAG CTGGAATCTC CTTTGCCATC CCTTCTGATC GTCTTCGAGA GTTCTGTCAT 1020  
 CGTGGGGAAG AGAAGAATTC CTCCTCCGGA ATCAGTGGGT CCCAGCGGCG CTACATTGGG 1080  
 GTGATGATGC TGACCCTGAG TCCAGGGGCT GGTCTGCGGC CTGGTGATGT GATTTTGGCC 1140  
 ATTGGGGAGC AGATGTGACA AAATGCTGAA GATGTTTATG AAGCTGTTTC AACCCAATCC 1200  
 CAGTTGGCAG TGCAGATCCG GCGGGGACGA GAAACACTGA CCTTATATGT GACCCCTGAG 1260  
 GTCACAGAA GAATAGATCA CCAAGAGTAT GAGGCTCCTG CTCTGATTTC CTCCTGCTCT 1320  
 TTCTGGCTGA GGTTCCTGAGG GCACCGAGAC AGAGGGTTAA ATGAACCACT GGGGGCAGGT 1380  
 CCCTCCAACC ACCAGCACTG ACTCCTGGGC TCTGAAGAAT CACAGAAACA CTTTTTATAT 1440  
 AAAATAAAAT TATACCTAGC AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 1500  
 AAA 1503

【0076】

配列番号：6

Leu Arg Ala Trp Arg Ala Leu Gly Gly Ile Arg Trp Gly Arg Arg Pro  
 1 5 10 15  
 Arg Leu Thr Pro Asp Leu Arg Ala Leu Leu Thr Ser Gly Thr Ser Asp  
 20 25 30  
 Pro Arg Ala Arg Val Thr Tyr Gly Thr Pro Ser Leu Trp Ala Arg Leu  
 35 40 45  
 Ser Val Gly Val Thr Glu Pro Arg Ala Cys Leu Thr Ser Gly Thr Pro  
 50 55 60  
 Gly Pro Arg Ala Gln Leu Thr Ala Val Thr Pro Asp Thr Arg Thr Arg  
 65 70 75 80  
 Glu Ala Ser Glu Asn Ser Gly Thr Arg Ser Arg Ala Trp Leu Ala Val  
 85 90 95  
 Ala Leu Gly Ala Gly Gly Ala Val Leu Leu Leu Trp Gly Gly Gly  
 100 105 110  
 Arg Gly Pro Pro Ala Val Leu Ala Ala Val Pro Ser P

39 40  
 Met Val Gln Asn Ala Glu Asp Val Tyr Glu Ala Val Arg Thr Gln Ser  
 385 390 395 400  
 Gln Leu Ala Val Gln Ile Arg Arg Gly Arg Glu Thr Leu Thr Leu Tyr  
 405 410 415  
 Val Thr Pro Glu Val Thr Glu  
 420

【0077】

配列番号: 7

GGCCGGAAGG GCTAGCGGTC CCAGCATACC CCGCGGCCCC TTGGGCGGTC TCACAACTCG 60  
 CGTCCGGCGG AGACCACAAT TCCCGGCATT CGTGGGGCAT GGAGGAGTCG GCCTCCCGGA 120  
 ATCCTGGTCC CGGCGTGAC TTCTGAAGGA CTCAGGTAC CGGCGTGCCC CGCGTCCTAC 180  
 TGTCCGCTG CTCGCGTCCT GGGTGCCGCC TCTGAGTAGG GCGGGCGAGG AGGCAGCCAA 240  
 GGCGGAGCTG ATG GCT GCG CCG AGG GCG GGG CGG GGT GCA GGC TGG AGC 289  
 Met Ala Ala Pro Arg Ala Gly Arg Gly Ala Gly Trp Ser  
 1 5 10  
 CTT CGG GCA TGG CGG GCT TTG GGG GGC ATT TGC TGG GGG AGG AGA CCC 337  
 Leu Arg Ala Trp Arg Ala Leu Gly Gly Ile Cys Trp Gly Arg Arg Pro  
 15 20 25  
 CGT TTG ACC CCT GAC CTC CGG GCC CTG CTG ACG TCA GGA ACT TCT GAC 385  
 Arg Leu Thr Pro Asp Leu Arg Ala Leu Leu Thr Ser Gly Thr Ser Asp  
 30 35 40 45  
 CCC CGG GCC CGA GTG ACT TAT GGG ACC CCC AGT CTC TGG GCC CGG TTG 433  
 Pro Arg Ala Arg Val Thr Tyr Gly Thr Pro Ser Leu Trp Ala Arg Leu  
 50 55 60  
 TCT GTT GGG GTC ACT GAA CCC CGA GCA TGC CTG ACG TCT GGG ACC CCG 481  
 Ser Val Gly Val Thr Glu Pro Arg Ala Cys Leu Thr Ser Gly Thr Pro  
 65 70 75  
 GGT CCC CGG GCA CAA CTG ACT GCG GTG ACC CCA GAT ACC AGG ACC CGG 529  
 Gly Pro Arg Ala Gln Leu Thr Ala Val Thr Pro Asp Thr Arg Thr Arg  
 80 85 90  
 GAG GCC TCA GAG AAC TCT GGA ACC CGT TCG CGC GCG TGG CTG GCG GTG 577  
 Glu Ala Ser Glu Asn Ser Gly Thr Arg Ser Arg Ala Trp Leu Ala Val  
 95 100 105  
 GCG CTG GGC GCT GGG GGG GCA GTG CTG TTG TTG TTG TGG GGC GGG GGT 625  
 Ala Leu Gly Ala Gly Gly Ala Val Leu Leu Leu Leu Trp Gly Gly Gly  
 110 115 120 125  
 CGG GGT CCT CCG GCC GTC CTC GCC GCC GTC CCT AGC CCG CCG CCC GCT 673  
 Arg Gly Pro Pro Ala Val Leu Ala Ala Val Pro Ser Pro Pro Pro Ala  
 130 135 140  
 TCT CCC CGG AGT CAG TAC AAC TTC ATC GCA GAT GTG GTG GAG AAG ACA 721  
 Ser Pro Arg Ser Gln Tyr Asn Phe Ile Ala Asp Val Val Glu Lys Thr  
 145 150 155  
 GCA CCT GCC GTG GTC TAT ATC GAG ATC CTG GAC CGG CAC CCT TTC TTG 769  
 Ala Pro Ala Val Val Tyr Ile Glu Ile Leu Asp Arg His Pro Phe Leu  
 160 165 170  
 GGC CGC GAG GTC CCT ATC TCG AAC GGC TCA GGA TTC GTG GTG GCT GCC 817  
 Gly Arg Glu Val Pro Ile Ser Asn Gly Ser Gly Phe Val Val Ala Ala  
 175 180 185  
 GAT GGG CTC ATT GTC ACC AAC GCC CAT GTG GTG GCT GAT CCG CGC AGA 865  
 Asp Gly Leu Ile Val Thr Asn Ala His Val Val Ala Asp Arg Arg Arg

41		42	
190	195	200	205
GTC CGT GTG AGA CTG CTA AGC GGC GAC ACG TAT GAG GCC GTG GTC ACA			913
Val Arg Val Arg Leu Leu Ser Gly Asp Thr Tyr Glu Ala Val Val Thr			
	210	215	220
GCT GTG GAT CCC GTG GCA GAC ATC GCA ACG CTG AGG ATT CAG ACT AAG			961
Ala Val Asp Pro Val Ala Asp Ile Ala Thr Leu Arg Ile Gln Thr Lys			
	225	230	235
GAG CCT CTC CCC ACG CTG CCT CTG GGA CGC TCA GCT GAT GTC CGG CAA			1009
Glu Pro Leu Pro Thr Leu Pro Leu Gly Arg Ser Ala Asp Val Arg Gln			
	240	245	250
GGG GAG TTT GTT GTT GCC ATG GGA AGT CCC TTT GCA CTG CAG AAC ACG			1057
Gly Glu Phe Val Val Ala Met Gly Ser Pro Phe Ala Leu Gln Asn Thr			
	255	260	265
ATC ACA TCC GGC ATT GTT AGC TCT GCT CAG CGT CCA GCC AGA GAC CTG			1105
Ile Thr Ser Gly Ile Val Ser Ser Ala Gln Arg Pro Ala Arg Asp Leu			
	270	275	280
GGA CTC CCC CAA ACC AAT GTG GAA TAC ATT CAA ACT GAT GCA GCT ATT			1153
Gly Leu Pro Gln Thr Asn Val Glu Tyr Ile Gln Thr Asp Ala Ala Ile			
	290	295	300
GAT TTT GGA AAC TCT GGA GGT CCC CTG GTT AAC CTG GAT GGG GAG GTG			1201
Asp Phe Gly Asn Ser Gly Gly Pro Leu Val Asn Leu Asp Gly Glu Val			
	305	310	315
ATT GGA GTG AAC ACC ATG AAG GTC ACA GCT GGA ATC TCC TTT GCC ATC			1249
Ile Gly Val Asn Thr Met Lys Val Thr Ala Gly Ile Ser Phe Ala Ile			
	320	325	330
CCT TCT GAT CGT CTT CGA GAG TTT CTG CAT CGT GGG GAA AAG AAG AAT			1297
Pro Ser Asp Arg Leu Arg Glu Phe Leu His Arg Gly Glu Lys Lys Asn			
	335	340	345
TCC TCC TCC GGA ATC AGT GGG TCC CAG CGG CGC TAC ATT GGG GTG ATG			1345
Ser Ser Ser Gly Ile Ser Gly Ser Gln Arg Arg Tyr Ile Gly Val Met			
	350	355	360
ATG CTG ACC CTG AGT CCC AGC ATC CTT GCT GAA CTA CAG CTT CGA GAA			1393
Met Leu Thr Leu Ser Pro Ser Ile Leu Ala Glu Leu Gln Leu Arg Glu			
	370	375	380
CCA AGC TTT CCC GAT GTT CAG CAT GGT GTA CTC ATC CAT AAA GTC ATC			1441
Pro Ser Phe Pro Asp Val Gln His Gly Val Leu Ile His Lys Val Ile			
	385	390	395
CTG GGC TCC CCT GCA CAC CGG GCT GGT CTG CGG CCT GGT GAT GTG ATT			1489
Leu Gly Ser Pro Ala His Arg Ala Gly Leu Arg Pro Gly Asp Val Ile			
	400	405	410
TTG GCC ATT GGG GAG CAG ATG GTA CAA AAT GCT GAA GAT GTT TAT GAA			1537
Leu Ala Ile Gly Glu Gln Met Val Gln Asn Ala Glu Asp Val Tyr Glu			
	415	420	425
GCT GTT CGA ACC CAA TCC CAG TTG GCA GTG CAG ATC CGG CGG GGA CGA			1585
Ala Val Arg Thr Gln Ser Gln Leu Ala Val Gln Ile Arg Arg Gly Arg			
	430	435	440
GAA ACA CTG ACC TTA TAT GTG ACC CCT GAG GTC ACA GAA TGAATAGATC ACC			1637
Glu Thr Leu Thr Leu Tyr Val Thr Pro Glu Val Thr Glu			
	450	455	
AAGAGTATGA GGCTCCTGCT CTGATTCCT CCTTGCTTT CTGGCTGAGG TTCTGAGGGC			1697



43

44

ACCGAGACAG AGGGTTAAAT GAACCACTGG GGGCAGGTCC CTCCAACCAC CAGCACTGAC 1757  
 TCCTGGGCTC TGAAGAATCA CAGAAACACT TTTTATATAA AATAAAATTA TACCTAGCAA 1817  
 CATAAAAAAA AAAAAAAA 1835

【0078】

配列番号: 8

Met Ala Ala Pro Arg Ala Gly Arg Gly Ala Gly Trp Ser Leu Arg Ala  
 1 5 10 15  
 Trp Arg Ala Leu Gly Gly Ile Cys Trp Gly Arg Arg Pro Arg Leu Thr  
 20 25 30  
 Pro Asp Leu Arg Ala Leu Leu Thr Ser Gly Thr Ser Asp Pro Arg Ala  
 35 40 45  
 Arg Val Thr Tyr Gly Thr Pro Ser Leu Trp Ala Arg Leu Ser Val Gly  
 50 55 60  
 Val Thr Glu Pro Arg Ala Cys Leu Thr Ser Gly Thr Pro Gly Pro Arg  
 65 70 75 80  
 Ala Gln Leu Thr Ala Val Thr Pro Asp Thr Arg Thr Arg Glu Ala Ser  
 85 90 95  
 Glu Asn Ser Gly Thr Arg Ser Arg Ala Trp Leu Ala Val Ala Leu Gly  
 100 105 110  
 Ala Gly Gly Ala Val Leu Leu Leu Trp Gly Gly Gly Arg Gly Pro  
 115 120 125  
 Pro Ala Val Leu Ala Ala Val Pro Ser Pro Pro Pro Ala Ser Pro Arg  
 130 135 140  
 Ser Gln Tyr Asn Phe Ile Ala Asp Val Val Glu Lys Thr Ala Pro Ala  
 145 150 155 160  
 Val Val Tyr Ile Glu Ile Leu Asp Arg His Pro Phe Leu Gly Arg Glu  
 165 170 175  
 Val Pro Ile Ser Asn Gly Ser Gly Phe Val Val Ala Ala Asp Gly Leu  
 180 185 190  
 Ile Val Thr Asn Ala His Val Val Ala Asp Arg Arg Arg Val Arg Val  
 195 200 205  
 Arg Leu Leu Ser Gly Asp Thr Tyr Glu Ala Val Val Thr Ala Val Asp  
 210 215 220  
 Pro Val Ala Asp Ile Ala Thr Leu Arg Ile Gln Thr Lys Glu Pro Leu  
 225 230 235 240  
 Pro Thr Leu Pro Leu Gly Arg Ser Ala Asp Val Arg Gln Gly Glu Phe  
 245 250 255  
 Val Val Ala Met Gly Ser Pro Phe Ala Leu Gln Asn Thr Ile Thr Ser  
 260 265 270  
 Gly Ile Val Ser Ser Ala Gln Arg Pro Ala Arg Asp Leu Gly Leu Pro  
 275 280 285  
 Gln Thr Asn Val Glu Tyr Ile Gln Thr Asp Ala Ala Ile Asp Phe Gly  
 290 295 300  
 Asn Ser Gly Gly Pro Leu Val Asn Leu Asp Gly Glu Val Ile Gly Val  
 305 310 315 320  
 Asn Thr Met Lys Val Thr Ala Gly Ile Ser Phe Ala Ile Pro Ser Asp  
 325 330 335  
 Arg Leu Arg Glu Phe Leu His Arg Gly Glu Lys Lys Asn Ser Ser Ser  
 340 345 350  
 Gly Ile Ser Gly Ser Gln Arg Arg Tyr Ile Gly Val Met Met Leu Thr

【0079】

TGGGACAGCG	AGCTCCGGGG	TCCGCGGTTT	CACATCGGAA	ACAAAACAGC	GGCTGGTCTG	60
GAAGGAACCT	GAGCTACGAG	CCGCGGCGGC	AGCGGGCGCG	CGGGGAAGCG	TATACCTAAT	120
CTGGGAGCCT	GCAAGTGACA	ACAGCCTTTG	CGGTCCTTAG	ACAGCTTGGC	CTGGAGGAGA	180
ACACATGAAA	GAAGAACCT	CAAGAGGCTT	TGTTTTCTGT	GAAACAGTAT	TTCTATACAG	240
TTGCTCCAAT	GACAGAGTTA	CCTGCACCGT	TGTCCTACTT	CCAGAATGCA	CAGATGTCTG	300
AGGACAACCA	CCTGAGCAAT	ACTGTACGTA	GCCAGAATGA	CAATAGAGAA	CGGCAGGAGC	360
ACAACGACAG	ACGGAGCCTT	GGCCACCCTG	AGCCATTATC	TAATGGACGA	CCCCAGGGTA	420
ACTCCCGGCA	GGTGGTGGAG	CAAGATGAGG	AAGAAGATGA	GGAGCTGACA	TTGAAATATG	480
GCGCCAAGCA	TGTGATCATG	CTCTTTGTCC	CTGTGACTCT	CTGCATGGTG	GTGGTCGTGG	540
CTACCATTAA	GTCAGTCAGC	TTTTATACCC	GGAAGGATGG	GCAGCTAATC	TATACCCCAT	600
TCACAGAAGA	TACCGAGACT	GTGGGCCAGA	GAGCCCTGCA	CTCAATTCTG	AATGCTGCCA	660
TCATGATCAG	TGTCATTGTT	GTCATGACTA	TCCTCCTGGT	GGTTCGTGAT	AAATACAGGT	720
GCTATAAGGT	CATCCATGCC	TGGCTTATTA	TATCATCTCT	ATTGTTGCTG	TTCTTTTTTT	780
CATTCAATTA	CTTGGGGGAA	GTGTTTAAAA	CCTATAACGT	TGCTGTGGAC	TACATTACTG	840
TTGCACTCCT	GATCTGGAAT	TTTGGTGTGG	TGGGAATGAT	TTCCATTAC	TGAAAAGGTC	900
CACCTCGACT	CCAGCAGGCA	TATCTCATT	TGATTAGTGC	CCTCATGGCC	CTGGTGTTTA	960
TCAAGTACCT	CCCTGAATGG	ACTGCGTGGC	TCATCTTGGC	TGTGATTTC	GTATATGATT	1020
TAGTGGCTGT	TTTGTGTCCG	AAAGGTCCAC	TTCTGATGCT	GGTTGAAACA	GCTCAGGAGA	1080
GAAATGAAAC	GCTTTTTCCA	GCTCTCATTT	ACTCCTCAAC	AATGGTGTGG	TTGGTGAATA	1140
TGGCAGAAGG	AGACCCGGAA	GCTCAAAGGA	GAGTATCCAA	AAATTCCAAG	TATAATGCAG	1200
AAAGCACAGA	AAGGGAGTCA	CAAGACACTG	TTGCAGAGAA	TGATGATGGC	GGGTTCAGTG	1260
AGGAATGGGA	AGCCCAGAGG	GACAGTCATC	TAGGGCCTCA	TCGCTCTACA	CCTGAGTCAC	1320
GAGCTGCTGT	CCAGGAACCT	TCCAGCAGTA	TCCTCGCTGG	TGAAGACCCA	GAGGAAAGGG	1380
GAGTAAACT	TGGATTGGGA	GATTTCAAT	TCTACAGTGT	TCTGGTTGGT	AAAGCCTCAG	1440
CAACAGCCAG	TGGAGACTGG	AACACAACCA	TAGCCTGTTT	CGTAGCCATA	TTAATTGTTT	1500
TGTGCCTTAC	ATTATTACTC	CTTGCCATTT	TCAAGAAAGC	ATTGCCAGCT	CTTCCAATCT	1560
CCATCACCTT	TGGGCTTGTT	TTCTACTTTG	CCACAGATTA	TCTTGTACAG	CCTTTTATGG	1620
ACCAATTAGC	ATTCCATCAA	TTTTATATCT	AGCATATTTG	CGTTTAGAAT	CCCATGGATG	1680
TTTCTTCTTT	GACTATAACC	AAATCTGGGG	AGGACAAAGG	TGATTTTCCT	GTGTCCACAT	1740
CTAACAAAGT	CAAGATTCCC	GGCTGGACTT	TTGCAGCTTC	CTTCCAAGTC	TTCTGACCA	1800
CCTTGCACTA	TTGGACTTTG	GAAGGAGGTG	CCTATAGAAA	ACGATTTTGA	ACATACTTCA	1860
TCGCAGTGGA	CTGTGTCCTT	CGGTGCAGAA	ACTACCAGAT	TTGAGGGAGC	AGGTCAAGGA	1920
GATATGATAG	GCCCCGAAGT	TGCTGTGCCC	CATCAGCAGC	TTGACGCGTG	GTACAGGAC	1980
GATTTCACTG	ACACTGCGAA	CTCTCAGGAC	TACCGGTTAC	CAAGAGGTTA	GGTGAAGTGG	2040
TTTAAACCAA	ACGGAACTCT	TCATCTTAAA	CTACACGTTG	AAAATCAACC	CAATAATTCT	2100

47

48

GTATTAACG AATTCTGAAC TTTTCAGGAG GTACTGTGAG GAAGAGCAGG CACCAGCAGC 2160  
 AGAATGGGGA ATGGAGAGGT GGGCAGGGGT TCCAGCTTCC CTTTGATTTT TTGCTGCAGA 2220  
 CTCATCCTTT TTAATGAGA CTGTGTTTCC CCTCTCTTG AGTCAAGTCA AATATGTAGA 2280  
 TTGCCTTTGG CAATTCTTCT TCTCAAGCAC TGACACTCAT TACCGTCTGT GATTGCCATT 2340  
 TCTTCCAAG GCCAGTCTGA ACCTGAGGTT GCTTTATCCT AAAAGTTTAA ACCTCAGGTT 2400  
 CCAAATTCAG TAAATTTTGG AAACAGTACA GCTATTCTC ATCAATTCTC TATCATGTTG 2460  
 AAGTCAAATT TGGATTTTCC ACCAAATTCT GAATTTGTAG ACATACTTGT ACGCTCACTT 2520  
 GCCCCAGAT GCCTCCTCTG TCCTCATTCT TCTCTCCAC ACAAGCAGTC TTTTCTACA 2580  
 GCCAGTAAGG CAGCTCTGTC RTGGTAGCAG ATGGTCCCAT TATTCTAGGG TCTTACTCTT 2640  
 TGTATGATGA AAAGAATGTG TTATGAATCG GTGCTGTCAG CCCTGCTGTC AGACCTTCTT 2700  
 CCACAGCAAA TGAGATGTAT GCCCAAAGCG GTAGAATTAA AGAAGAGTAA AATGGCTGTT 2760  
 GAAG 2764

[0080]

配列番号: 10

Met Thr Glu Leu Pro Ala Pro Leu Ser Tyr Phe Gln Asn Ala Gln Met  
 1 5 10 15  
 Ser Glu Asp Asn His Leu Ser Asn Thr Val Arg Ser Gln Asn Asp Asn  
 20 25 30  
 Arg Glu Arg Gln Glu His Asn Asp Arg Arg Ser Leu Gly His Pro Glu  
 35 40 45  
 Pro Leu Ser Asn Gly Arg Pro Gln Gly Asn Ser Arg Gln Val Val Glu  
 50 55 60  
 Gln Asp Glu Glu Glu Asp Glu Glu Leu Thr Leu Lys Tyr Gly Ala Lys  
 65 70 75 80  
 His Val Ile Met Leu Phe Val Pro Val Thr Leu Cys Met Val Val Val  
 85 90 95  
 Val Ala Thr Ile Lys Ser Val Ser Phe Tyr Thr Arg Lys Asp Gly Gln  
 100 105 110  
 Leu Ile Tyr Thr Pro Phe Thr Glu Asp Thr Glu Thr Val Gly Gln Arg  
 115 120 125  
 Ala Leu His Ser Ile Leu Asn Ala Ala Ile Met Ile Ser Val Ile Val  
 130 135 140  
 Val Met Thr Ile Leu Leu Val Val Leu Tyr Lys Tyr Arg Cys Tyr Lys  
 145 150 155 160  
 Val Ile His Ala Trp Leu Ile Ile Ser Ser Leu Leu Leu Phe Phe  
 165 170 175  
 Phe Ser Phe Ile Tyr Leu Gly Glu Val Phe Lys Thr Tyr Asn Val Ala  
 180 185 190  
 Val Asp Tyr Ile Thr Val Ala Leu Leu Ile Trp Asn Phe Gly Val Val  
 195 200 205  
 Gly Met Ile Ser Ile His Trp Lys Gly Pro Leu Arg Leu Gln Gln Ala  
 210 215 220  
 Tyr Leu Ile Met Ile Ser Ala Leu Met Ala Leu Val Phe Ile Lys Tyr  
 225 230 235 240  
 Leu Pro Glu Trp Thr Ala Trp Leu Ile Leu Ala Val Ile Ser Val Tyr  
 245 250 255  
 Asp Leu Val Ala Val Leu Cys Pro Lys Gly Pro Leu Arg Met Leu Val  
 260 265 270  
 Glu Thr Ala Gln Glu Arg Asn Glu Thr Leu Phe Pro Ala Leu Ile Tyr  
 275 280 285

49 50  
 Ser Ser Thr Met Val Trp Leu Val Asn Met Ala Glu Gly Asp Pro Glu  
 290 295 300  
 Ala Gln Arg Arg Val Ser Lys Asn Ser Lys Tyr Asn Ala Glu Ser Thr  
 305 310 315 320  
 Glu Arg Glu Ser Gln Asp Thr Val Ala Glu Asn Asp Asp Gly Gly Phe  
 325 330 335  
 Ser Glu Glu Trp Glu Ala Gln Arg Asp Ser His Leu Gly Pro His Arg  
 340 345 350  
 Ser Thr Pro Glu Ser Arg Ala Ala Val Gln Glu Leu Ser Ser Ser Ile  
 355 360 365  
 Leu Ala Gly Glu Asp Pro Glu Glu Arg Gly Val Lys Leu Gly Leu Gly  
 370 375 380  
 Asp Phe Ile Phe Tyr Ser Val Leu Val Gly Lys Ala Ser Ala Thr Ala  
 385 390 395 400  
 Ser Gly Asp Trp Asn Thr Thr Ile Ala Cys Phe Val Ala Ile Leu Ile  
 405 410 415  
 Gly Leu Cys Leu Thr Leu Leu Leu Leu Ala Ile Phe Lys Lys Ala Leu  
 420 425 430  
 Pro Ala Leu Pro Ile Ser Ile Thr Phe Gly Leu Val Phe Tyr Phe Ala  
 435 440 445  
 Thr Asp Tyr Leu Val Gln Pro Phe Met Asp Gln Leu Ala Phe His Gln  
 450 455 460  
 Phe Tyr Ile  
 465

【0081】

配列番号：1 1

CGGAATTCCG TATGCTGGTT GAAACA

26

【0082】

配列番号：1 2

CGGGATCCTC AGGCTACGAA ACAGGCTAT

29

【0083】

配列番号：1 3

TATATCAGCG GTATGACCGA CCTCTATGCG TGGGATGAAT ACCGACGTCT GATGGCCGTA 60  
 GAACAATAAC CAGGCTTTTG TAAAGACGAA CAATAAATTT TTACCTTTTG CAGAACTTT 120  
 AGTTCGGAAC TTCAGGCTAT AAAACGAATC TGAAGAACAC AGCAATTTTG CGTTATCTGT 180  
 TAATCGAGAC TGAATACAT GAAAAAACC ACATTAGCAC TGAGTCGACT GGCTCTGAGT 240  
 TTAGGTTTGG CGTTATCTCC GCTCTCTGCA ACGGCGGCTG AGACTTCTTC AGCAACGACA 300  
 GCCCAGCAGA TGCCAAGCCT TGCACCGATG CTCGAAAAGG TGATGCCTTC AGTGGTCAGC 360  
 ATTAACGTAG AAGGTAGCAC AACCGTTAAT ACGCCGCGTA TGCCGCGTAA TTTCCAGCAG 420  
 TTCTTCGGTG ATGATTCTCC GTTCTGCCAG GAAGGTTCTC CGTTCCAGAG CTCTCCGTTT 480  
 TGCCAGGGTG GCCAGGGCGG TAATGGTGGC GGCCAGCAAC AGAAATTCAT GCGCGTGGGT 540  
 TCCGCGCTCA TCATTGATGC CGATAAAGGC TATGTCGTCA CCAACAACCA CGTTGTTGAT 600  
 AACGCGACGG TCATTAAAGT TCAACTGAGC GATGGCCGTA AGTTCGACGC GAAGATGGTT 660  
 GGCAAAGATC CCGGCTCTGA TATCGCGCTG ATCCAAATCC AGAACCCGAA AAACCTGACC 720  
 GCAATTAAGA TGGCGGATTC TGATGCACTG CGCGTGGGTG ATTACACCGT AGGGATTGGT 780  
 AACCCGTTTG GTCTGGGCGA GACGGTAACT TCCGGGATTG TCTCTGCGCT GGGGCGTAGC 840  
 GGCTGAATG CCGAAACTA CGAAACTTC ATCCAGACCG ATGCAGCGAT CAACCGTGGT 900  
 AACTCCGGTG GTGCGCTGGT TAACCTGAAC GGCGAACTGA TCGGTATCAA CACCGCGATC 960  
 CTCGCCACCGG ACGGCGGCAA CATCGGTATC GGTTTTGCTA TCCCGAGTAA CATGGTGAAG 1020  
 AACCTGACCT CGCAGATGGT GGAATACGGC CAGGTGAAAC GCGGTGAGCT GGGTATTATG 1080

51

52

GGGACTGAGC TGAAGTCCGA ACTGGCGAAA GCGATGAAAG TTGACGCCCA GCGCGGTGCT 1140  
 TTCGTAAGCC AGGTTCCTGCC TAATTCCTCC GCTGCAAAAG CGGGCATTAA AGCGGGTGAT 1200  
 GTGATCACCT CACTGAACGG TAAGCCGATC AGCAGCTTTG CCGCACTGCG TGCTCAGGTG 1260  
 GGTACTATGC CGGTAGGCAG CAAACTGACC CTGGGCTTAC TGGCGGACGG TAAGCAGGTT 1320  
 AACGTGAACC TGGAACTGCA GCAGAGCAGC CAGAATCAGG TTGATTCCAG CTCCATCTTC 1380  
 AACGGCATTG AAGGCGCTGA GATGAGCAAC AAAGGCAAAG ATCAGGGCGT GGTAGTGAAC 1440  
 AACGTGAAAA CGGGCACTCC GGCTGCGCAG ATCGGCCTGA AGAAAGGTGA TGTGATTATT 1500  
 GGCGCGAACC AGCAGGCAGT GAAAAACATC GCTGAACTGC GTAAAGTTCT CGACAGCAAA 1560  
 CCGTCTGTGC TGGCACTCAA CATTACGGCG GCGACCGCC ATCTACCTGT TAATGCAGTA 1620  
 ATCTCCCTCA ACCCCTTCCT GAAAACGGGA AGGGGTCTC CTTACAATCT GTGAACTCA 1680  
 CCACAACCTC ATACATCTTC ATCATCTTT AGGCATTGCG ACAATGCCGT ACGTTACGTA 1740  
 CTTCTTATG CTAAGCCGTG CATAACGGAG GACTTATGGC TGGCTGGCAT CTTGATACCA 1800  
 AAATGGCGCA GGATATCGTG GCACGTACCA TCGCATCAT CGATACCAAT ATCA 1854

[0084]

配列番号: 1 4

Met Lys Lys Thr Thr Leu Ala Leu Ser Arg Leu Ala Leu Ser Leu Gly  
 1 5 10 15  
 Leu Ala Leu Ser Pro Leu Ser Ala Thr Ala Ala Glu Thr Ser Ser Ala  
 20 25 30  
 Thr Thr Ala Gln Gln Met Pro Ser Leu Ala Pro Met Leu Glu Lys Val  
 35 40 45  
 Met Pro Ser Val Val Ser Ile Asn Val Glu Gly Ser Thr Thr Val Asn  
 50 55 60  
 Thr Pro Arg Met Pro Arg Asn Phe Gln Gln Phe Phe Gly Asp Asp Ser  
 65 70 75 80  
 Pro Phe Cys Gln Glu Gly Ser Pro Phe Gln Ser Ser Pro Phe Cys Gln  
 85 90 95  
 Gly Gly Gln Gly Gly Asn Gly Gly Gly Gln Gln Gln Lys Phe Met Ala  
 100 105 110  
 Leu Gly Ser Gly Val Ile Ile Asp Ala Asp Lys Gly Tyr Val Val Thr  
 115 120 125  
 Asn Asn His Val Val Asp Asn Ala Thr Val Ile Lys Val Gln Leu Ser  
 130 135 140  
 Asp Gly Arg Lys Phe Asp Ala Lys Met Val Gly Lys Asp Pro Arg Ser  
 145 150 155 160  
 Asp Ile Ala Leu Ile Gln Ile Gln Asn Pro Lys Asn Leu Thr Ala Ile  
 165 170 175  
 Lys Met Ala Asp Ser Asp Ala Leu Arg Val Gly Asp Tyr Thr Val Gly  
 180 185 190  
 Ile Gly Asn Pro Phe Gly Leu Gly Glu Thr Val Thr Ser Gly Ile Val  
 195 200 205  
 Ser Ala Leu Gly Arg Ser Gly Leu Asn Ala Glu Asn Tyr Glu Asn Phe  
 210 215 220  
 Ile Gln Thr Asp Ala Ala Ile Asn Arg Gly Asn Ser Gly Gly Ala Leu  
 225 230 235 240  
 Val Asn Leu Asn Gly Glu Leu Ile Gly Ile Asn Thr Ala Ile Leu Ala  
 245 250 255  
 Pro Asp Gly Gly Asn Ile Gly Ile Gly Phe Ala Ile Pro Ser Asn Met  
 260 265 270  
 Val Lys Asn Leu Thr Ser Gln Met Val Glu Tyr Gly Gln Val Lys Arg

【0085】

24

25

CCGGCCCTCG	CCCTGTCCGC	CGCCACCGCC	GCCGCCGCCA	GAGTCGCCAT	GCAGATCCCG	60
CGCGCCGCTC	TTCTCCCGCT	GCTGCTGCTG	CTGCTGGCGG	CGCCCGCCTC	GGCGCAGCTG	120
TCCCGGGCCG	GCCGCTCGGC	GCCTTTGGCC	GCCGGGTGCC	CAGACCGCTG	CGAGCCGGCG	180
CGTGCCCGC	CGCAGCCGGA	GCACTGCGAG	GGCGGCCGGG	CCCGGGACGC	GTGGGGCTGC	240
TGCGAGGTGT	GCGGCGCGCC	CGAGGGCGCC	GCGTGCGGCC	TGCAGGAGGG	CCCGTGCGGC	300
GAGGGGCTGC	AGTGCCTGGT	GCCCTTCGGG	GTGCCAGCCT	CGGCCACGGT	GCGGCGGGCG	360
GCGCAGGCGC	GCCTCTGTGT	GTGCGCCAGC	AGCGAGCCGG	TGTGCGGCAG	CGACGCCAAC	420
ACCTACGCCA	ACCTGTGCCA	GCTGCGCGCC	GCCAGCCGCC	GCTCCGAGAG	GCTGCACCGG	480
CGCCCGGTCA	TCGTCTGCA	GCGCGGAGCC	TGCGGCCAAG	GGCAGGAAGA	TCCAACACT	540
TTGCGCCATA	AATATAACTT	TATCGCGGAC	GTGGTGAGAG	AGATCGCCCC	TGCCGTGGTT	600
CATATCGAAT	TGTTTCGCAA	GCTTCCGTTT	TCTAAACGAG	AGGTGCCCGT	GGCTAGTGGG	660
TCTGGGTTTA	TTGTGTCCGA	AGATGGACTG	ATCGTGACAA	ATGCCACAGT	GGTGACCAAC	720
AAGCACCGGG	TCAAAGTTGA	GCTGAAGAAC	GGTGCCACTT	ACGAAGCCAA	AATCAAGGAT	780
GTGGATGAGA	AAGCAGACAT	CGCACTCATC	AAAATTGACC	ACCAGGGCAA	GCTGCCTGTC	840
CTGCTGCTTG	GCCGCTCCTC	AGAGCTGCGG	CCGGGAGAGT	TCGTGGTCGC	CATCGGAAGC	900
CCGTTTTTCC	TTCAAAACAC	AGTCACCACC	GGGATCGTGA	GCACCACCCA	GCGAGGCGGC	960

55

56

AAAGAGCTGG GGCTCCGCAA CTCAGACATG GACTACATCC AGACCGACGC CATCATCAAC 1020  
 TATGGAAACT CGGGAGGCCG GTTAGTAAAC CTGGACGGTG AAGTGATTGG AATTAACACT 1080  
 TTGAAAGTGA CAGCTGGAAT CTCCTTTGCA ATCCCATCTG ATAAGATTAA AAAGTTCCTC 1140  
 ACGGAGTCCC ATGACCGACA GGCCAAAGGA AAAGCCATCA CCAAGAAGAA GTATATTGGT 1200  
 ATCCGAATGA TGCTACTCAC GTCCAGCAAA GCCAAAGAGC TGAAGGACCG GCACCGGGAC 1260  
 TTCCAGACG TGATCTCAGG AGCGTATATA ATTGAAGTAA TTCCTGATAC CCCAGCAGAA 1320  
 GCTGKTGGTC TCAAGGAAAA CGACGTCATA ATCAGCATCA ATGGACAGTC CGTGGTCTCC 1380  
 GCCAATGATG TCAGCGACGT CATTAAAAGG GAAAGCACCC TGAACATGGT GGTCCGCGAG 1440  
 GGTAATGAAG ATATCATGAT CACAGTGATT CCCGAAGAAA TTGACCCATA GGCAGAGGCA 1500  
 TGAGCTGGAC TTCATGTTTC CCTCAAAGAC TCTCCCGTGG ATGACGGATG AGGACTCTGG 1560  
 GCTGCTGGAA TAGGACACTC AAGACTTTTG ACTGCCATTT TGTGTGTCA GTGGAGACTC 1620  
 CCTGGCCAAC AGAATCCTTC TTGATAGTTT GCAGGCAAAA CAAATGTAAT GTTGACAGATC 1680  
 CGCAGGCAGA AGCTCTGCCC TTCTGTATCC TATGTATGCA GTGTGCTTTT TCTTGCCAGC 1740  
 TTGGGCCATT CTTGCTTAGA CAGTCAGCAT TTGTCTCTC CTTTAACTGA GTCATCATCT 1800  
 TAGTCCAAC AATGCAGTCG ATACAATGCG TAGATAGAAG AAGCCCCACG GGAGCCAGGA 1860  
 TGGGACTGGT CGTGTGTTTG CTTTTCTCCA AGTCAGCACC CAAAGGTCAA TGCACAGAGA 1920  
 CCCCCGGTGG GTGACGCTG GCTTCTCAAA CGCCCGAAGT TGCCTCTTTT AGGAATCTCT 1980  
 TTGGAATTGG GAGCAGCATG ACTCTGAGTT TGAGCTATTA AAGTACTTCT TACAAA 2036

【0088】配列番号: 18

\* \*特徴-213Gly/Va1多型性

Met Gln Ile Pro Arg Ala Ala Leu Leu Pro Leu Leu Leu Leu Leu  
 1 5 10 15  
 Ala Ala Pro Ala Ser Ala Gln Leu Ser Arg Ala Gly Arg Ser Ala Pro  
 20 25 30  
 Leu Ala Ala Gly Cys Pro Asp Arg Cys Glu Pro Ala Arg Cys Pro Pro  
 35 40 45  
 Gln Pro Glu His Cys Glu Gly Arg Ala Arg Asp Ala Cys Gly Cys  
 50 55 60  
 Cys Glu Val Cys Gly Ala Pro Glu Gly Ala Ala Cys Gly Leu Gln Glu  
 65 70 75 80  
 Gly Pro Cys Gly Glu Gly Leu Gln Cys Val Val Pro Phe Gly Val Pro  
 85 90 95  
 Ala Ser Ala Thr Val Arg Arg Arg Ala Gln Ala Gly Leu Cys Val Cys  
 100 105 110  
 Ala Ser Ser Glu Pro Val Cys Gly Ser Asp Ala Asn Thr Tyr Ala Asn  
 115 120 125  
 Leu Cys Gln Leu Arg Ala Ala Ser Arg Arg Ser Glu Arg Leu His Arg  
 130 135 140  
 Pro Pro Val Ile Val Leu Gln Arg Gly Ala Cys Gly Gln Gly Gln Glu  
 145 150 155 160  
 Asp Pro Asn Ser Leu Arg His Lys Tyr Asn Phe Ile Ala Asp Val Val  
 165 170 175  
 Glu Lys Ile Ala Pro Ala Val Val His Ile Glu Leu Phe Arg Lys Leu  
 180 185 190  
 Pro Phe Ser Lys Arg Glu Val Pro Val Ala Ser Gly Ser Gly Phe Ile  
 195 200 205  
 Val Ser Glu Asp Xaa Leu Ile Val Thr Asn Ala His Val Val Thr Asn  
 210 215 220  
 Lys His Arg Val Lys Val Glu Leu Lys Asn Gly Ala Thr Tyr Glu Ala  
 225 230 235 240  
 Lys Ile Lys Asp Val Asp Glu Lys Ala Asp Ile Ala Leu Ile Lys Ile

57

58

245 250 255  
 Asp His Gln Gly Lys Leu Pro Val Leu Leu Leu Gly Arg Ser Ser Glu  
 260 265 270  
 Leu Arg Pro Gly Glu Phe Val Val Ala Ile Gly Ser Pro Phe Ser Leu  
 275 280 285  
 Gln Asn Thr Val Thr Thr Gly Ile Val Ser Thr Thr Gln Arg Gly Gly  
 290 295 300  
 Lys Glu Leu Gly Leu Arg Asn Ser Asp Met Asp Tyr Ile Gln Thr Asp  
 305 310 315 320  
 Ala Ile Ile Asn Tyr Gly Asn Ser Gly Gly Pro Leu Val Asn Leu Asp  
 325 330 335  
 Gly Glu Val Ile Gly Ile Asn Thr Leu Lys Val Thr Ala Gly Ile Ser  
 340 345 350  
 Phe Ala Ile Pro Ser Asp Lys Ile Lys Lys Phe Leu Thr Glu Ser His  
 355 360 365  
 Asp Arg Gln Ala Lys Gly Lys Ala Ile Thr Lys Lys Lys Tyr Ile Gly  
 370 375 380  
 Ile Arg Met Met Ser Leu Thr Ser Ser Lys Ala Lys Glu Leu Lys Asp  
 385 390 395 400  
 Arg His Arg Asp Phe Pro Asp Val Ile Ser Gly Ala Tyr Ile Ile Glu  
 405 410 415  
 Val Ile Pro Asp Thr Pro Ala Glu Ala Gly Gly Leu Lys Glu Asn Asp  
 420 425 430  
 Val Ile Ile Ser Ile Asn Gly Gln Ser Val Val Ser Ala Asn Asp Val  
 435 440 445  
 Ser Asp Val Ile Lys Arg Glu Ser Thr Leu Asn Met Val Val Arg Arg  
 450 455 460  
 Gly Asn Glu Asp Ile Met Ile Thr Val Ile Pro Glu Glu Ile Asp Pro  
 465 470 475 480

【0089】

30

配列番号：19

ATGCTGAACA TCGGAAAGC TTGGTTCTCG

30

【0090】

配列番号：20

CCAACAGACA ACCGGGCCA GAGACT

26

【0091】

配列番号：21

TGCCTCCTCG CCCGCCCTAC TCAGA

25

【0092】

配列番号：22

CGTGGATCCC GAGAAAGAGG CGCAGGACGA GGAGGCAGAA CCCGACTGGC GCGTAGAGCA 60  
 GCAGCAGGAG CAGTAGGAAG CAGTCACCCG GAAGCCTGGG GCGAGAGGC GAACTGGTCA 120  
 GCGCCCGAAG GCCGAGAGCA CGCGGGGATC GGTCTCTTCC CGCCGGGTCT CTTACCGGTG 180  
 CGAGTCAAAG AGCCGCTCCG GCCCGGGCCC TGAGGGAAGC TCCATAACTG CTGCTTCAGG 240  
 AGCGCCCCGC CGTCGCCGCC GCCGCCATTT TCGCGCCCGG CCGCAGGGGC TCTTGGGAAG 300  
 GCGGAGTCTT TGGGCATCCG CCCGGGGTGA GGGGACCCGA AGTCTGAGG CGCGCCGGAA 360  
 GGGCTAGCGG TCCCAGCATA CCCCGGGGCC CCTTGGGCCG TCTCACAACCT CGCGTCCGGC 420  
 GGAGACCACA ATTCCCGGCA TTCGTGGGCG AGGGAGGAGT CGGCCTCCCG GAATCCTGGT 480  
 CCCGGCGTGC ACTTCTGAAG GACTTCAGGT ACCGGCGTGC CCCGCGTCCT ACTGTCCGCC 540  
 TGCTCGCGTC CTGGGTGCCG CCTCTGAGTA GGGCGGGCGA GGAGGCA 587



【0093】

配列番号: 23

CGTGGATCCC GAGAAAGAGG CGCAGGACGA GGAGGCAGAA CCCGACTGGC GCGTAGAGCA 60  
 GCAGCACGAG CAGTAGGAAG CAGTCACCCG GAAGCCTGGG GCGGAGAGGC GAAGTGGTCA 120  
 GCGCGCGAAG GCCGAGAGCA CGCGGGGATC GGTCTCTTCC CGCCGGGTCT CTTACCGGTG 180  
 CGAGTCAAAG AGCCGCTCCG GCGCCGGCCC TGAGGGAAGC TCCATAACTG CTGCTTCAGG 240  
 AGCGCCCGGC CGTCGCCGCC GCGGCCATTT TCGCGCCCGG CCGCAGGGGC TCTTGGGAAG 300  
 GCGGAGTCTT TGGGCATCCG CCCGGGGTGA GGGGACCCGA AGTCCTGAGG CGCGCCGGAA 360  
 GGGCTAGCGG TCCCAGCATA CCCCGCGGCC CCTTGGGCCG TCTCACAACCT CGCGTCCGGC 420  
 GGAGACCACA ATTCCCGGCA TTCGTGGGGC AGGGAGGAGT CGGCCTCCCG GAATCCTGGT 480  
 CCCGGCGTGC ACTTCTGAAG GACTTCAGGT ACCGGCGTGC CCCGCGTCCT ACTGTCCGCC 540  
 TGCTCGCGTC CTGGGTGCCG CCTCTGAGTA GGGCGGGCGA GGAGGCAGCC AAGGCGGAGC 600  
 TG ATG GCT GCG CCG AGG GCG GGG CGG GGT GCA GGC TGG AGC CTT CGG 647  
 Met Ala Ala Pro Arg Ala Gly Arg Gly Ala Gly Trp Ser Leu Arg  
 1 5 10 15  
 GCA TGG CGG GCT TTG GGG GGC ATT TGC TGG GGG AGG AGA CCC CGT TTG 695  
 Ala Trp Arg Ala Leu Gly Gly Ile Cys Trp Gly Arg Arg Pro Arg Leu  
 20 25 30  
 ACC CCT GAC CTC CGG GCC CTG CTG ACG TCA GGA ACT TCT GAC CCC CGG 743  
 Thr Pro Asp Leu Arg Ala Leu Leu Thr Ser Gly Thr Ser Asp Pro Arg  
 35 40 45  
 GCC CGA GTG ACT TAT GGG ACC CCC AGT CTC TGG GCC CGG TTG TCT GTT 791  
 Ala Arg Val Thr Tyr Gly Thr Pro Ser Leu Trp Ala Arg Leu Ser Val  
 50 55 60  
 GGG GTC ACT GAA CCC CGA GCA TGC CTG ACG TCT GGG ACC CCG GGT CCC 839  
 Gly Val Thr Glu Pro Arg Ala Cys Leu Thr Ser Gly Thr Pro Gly Pro  
 65 70 75  
 CGG GCA CAA CTG ACT GCG GTG ACC CCA GAT ACC AGG ACC CGG GAG GCC 887  
 Arg Ala Gln Leu Thr Ala Val Thr Pro Asp Thr Arg Thr Arg Glu Ala  
 80 85 90 95  
 TCA GAG AAC TCT GGA ACC CGT TCG CGC GCG TGG CTG GCG GTG GCG CTG 935  
 Ser Glu Asn Ser Gly Thr Arg Ser Arg Ala Trp Leu Ala Val Ala Leu  
 100 105 110  
 GGC GCT GGG GGG GCA GTG CTG TTG TTG TTG TGG GGC GGG GGT CGG GGT 983  
 Gly Ala Gly Gly Ala Val Leu Leu Leu Leu Trp Gly Gly Gly Arg Gly  
 115 120 125  
 CCT CCG GCC GTC CTC GCC GCC GTC CCT AGC CCG CCG CCC GCT TCT CCC 1031  
 Pro Pro Ala Val Leu Ala Ala Val Pro Ser Pro Pro Pro Ala Ser Pro  
 130 135 140  
 CGG AGT CAG TAC AAC TTC ATC GCA GAT GTG GTG GAG AAG ACA GCA CCT 1079  
 Arg Ser Gln Tyr Asn Phe Ile Ala Asp Val Val Glu Lys Thr Ala Pro  
 145 150 155  
 GCC GTG GTC TAT ATC GAG ATC CTG GAC CGG CAC CCT TTC TTG GGC CGC 1127  
 Ala Val Val Tyr Ile Glu Ile Leu Asp Arg His Pro Phe Leu Gly Arg  
 160 165 170 175  
 GAG GTC CCT ATC TCG AAC GGC TCA GGA TTC GTG GTG GCT GCC GAT GGG 1175  
 Glu Val Pro Ile Ser Asn Gly Ser Gly Phe Val Val Ala Ala Asp Gly  
 180 185 190  
 CTC ATT GTC ACC AAC GCC CAT GTG GTG GCT GAT CGG CGC AGA GTC CGT 1223  
 Leu Ile Val Thr Asn Ala His Val Val Ala Asp Arg Arg Arg Val Arg

61

62

195	200	205	
GTG AGA CTG CTA AGC GGC GAC ACG TAT GAG GCC GTG GTC ACA GCT GTG			1271
Val Arg Leu Leu Ser Gly Asp Thr Tyr Glu Ala Val Val Thr Ala Val			
210	215	220	
GAT CCC GTG GCA GAC ATC GCA ACG CTG AGG ATT CAG ACT AAG GAG CCT			1319
Asp Pro Val Ala Asp Ile Ala Thr Leu Arg Ile Gln Thr Lys Glu Pro			
225	230	235	
CTC CCC ACG CTG CCT CTG GGA CGC TCA GCT GAT GTC CGG CAA GGG GAG			1367
Leu Pro Thr Leu Pro Leu Gly Arg Ser Ala Asp Val Arg Gln Gly Glu			
240	245	250	255
TTT GTT GTT GCC ATG GGA AGT CCC TTT GCA CTG CAG AAC ACG ATC ACA			1415
Phe Val Val Ala Met Gly Ser Pro Phe Ala Leu Gln Asn Thr Ile Thr			
260	265	270	
TCC GGC ATT GTT AGC TCT GCT CAG CGT CCA GCC AGA GAC CTG GGA CTC			1463
Ser Gly Ile Val Ser Ser Ala Gln Arg Pro Ala Arg Asp Leu Gly Leu			
275	280	285	
CCC CAA ACC AAT GTG GAA TAC ATT CAA ACT GAT GCA GCT ATT GAT TTT			1511
Pro Gln Thr Asn Val Glu Tyr Ile Gln Thr Asp Ala Ala Ile Asp Phe			
290	295	300	
GGA AAC TCT GGA GGT CCC CTG GTT AAC CTG GAT GGG GAG GTG ATT GGA			1559
Gly Asn Ser Gly Gly Pro Leu Val Asn Leu Asp Gly Glu Val Ile Gly			
305	310	315	
GTG AAC ACC ATG AAG GTC ACA GCT GGA ATC TCC TTT GCC ATC CCT TCT			1607
Val Asn Thr Met Lys Val Thr Ala Gly Ile Ser Phe Ala Ile Pro Ser			
320	325	330	335
GAT CGT CTT CGA GAG TTT CTG CAT CGT GGG GAA AAG AAG AAT TCC TCC			1655
Asp Arg Leu Arg Glu Phe Leu His Arg Gly Glu Lys Lys Asn Ser Ser			
340	345	350	
TCC GGA ATC AGT GGG TCC CAG CGG CGC TAC ATT GGG GTG ATG ATG CTG			1703
Ser Gly Ile Ser Gly Ser Gln Arg Arg Tyr Ile Gly Val Met Met Leu			
355	360	365	
ACC CTG AGT CCC AGC ATC CTT GCT GAA CTA CAG CTT CGA GAA CCA AGC			1751
Thr Leu Ser Pro Ser Ile Leu Ala Glu Leu Gln Leu Arg Glu Pro Ser			
370	375	380	
TTT CCC GAT GTT CAG CAT GGT GTA CTC ATC CAT AAA GTC ATC CTG GGC			1799
Phe Pro Asp Val Gln His Gly Val Leu Ile His Lys Val Ile Leu Gly			
385	390	395	
TCC CCT GCA CAC CGG GCT GGT CTG CGG CCT GGT GAT GTG ATT TTG GCC			1847
Ser Pro Ala His Arg Ala Gly Leu Arg Pro Gly Asp Val Ile Leu Ala			
400	405	410	415
ATT GGG GAG CAG ATG GTA CAA AAT GCT GAA GAT GTT TAT GAA GCT GTT			1895
Ile Gly Glu Gln Met Val Gln Asn Ala Glu Asp Val Tyr Glu Ala Val			
420	425	430	
CGA ACC CAA TCC CAG TTG GCA GTG CAG ATC CGG CGG GGA CGA GAA ACA			1943
Arg Thr Gln Ser Gln Leu Ala Val Gln Ile Arg Arg Gly Arg Glu Thr			
435	440	445	
CTG ACC TTA TAT GTG ACC CCT GAG GTC ACA GAA TGAATAGATC ACCAAGACTA			1996
Leu Thr Leu Tyr Val Thr Pro Glu Val Thr Glu			
450	455		
TGAGGCTCCT GCTCTGATTT CCTCCTTGCC TTTCTGGCTG AGGTTCTGAG GGCACCGAGA			2056

63

64

CAGAGGGTTA AATGAACCAG TGGGGGCAGG TCCCTCCAAC CACCAGCACT GACTCCTGGG 2116  
 CTCTGAAGAA TCACAGAAAC ACTTTTATA TAAAATAAAA TTATACCTAG CAACATAAAA 2176  
 AAAAAAAAAA A 2187

【0094】

配列番号: 24

CGTGGATCCC GAGAAAGAGG CGCAGGACGA GGAGGCAGAA CCCGACTGGC GCGTAGAGCA 60  
 GCAGCACGAG CAGTAGGAAG CAGTCACCCG GAAGCCTGGG GGCAGAGAGC GAAGTGGTCA 120  
 GGCGCCGAAG GCCGAGAGCA CGCGGGGATC GGTCTCTTCC CGCCGGGTCT CTTACCGGTG 180  
 CGAGTCAAAG AGCCGCTCCG GCCCGGCCCG TGAGGGAAGC TCCATAACTG CTGCTTCAGG 240  
 AGCGCCCGGC CGTCGCCGCC GCCGCCATTT TCGCGCCCGG CCGCAGGGGC TCTTGGAAG 300  
 GCGGAGTCTT TGGGCATCCG CCCGGGGTGA GGGGACCCGA AGTCCTGAGG CGCGCCGAA 360  
 GGGCTAGCGG TCCCAGCATA CCCCGCGGCC CCTTGGGCCG TCTCACAAC TCGCTCCGGC 420  
 GGAGACCACA ATTCCCGGCA TTCGTGGGGC AGGGAGGAGT CGGCCTCCCG GAATCCTGGT 480  
 CCCGGCGTGC ACTTCTGAAG GACTTCAGGT ACCGGCGTGC CCCGCGTCTT ACTGTCCGCC 540  
 TGCTCGCGTC CTGGGTGCCG CCTCTGAGTA GGGCGGGCGA GGAGGCAGCC AAGCGGAGC 600  
 TG ATG GCT GCG CCG AGG GCG GGG CGG GGT GCA GGC TGG AGC CTT CGG 647

Met Ala Ala Pro Arg Ala Gly Arg Gly Ala Gly Trp Ser Leu Arg

1 5 10 15  
 GCA TGG CGG GCT TTG GGG GGC ATT CGC TGG GGG AGG AGA CCC CGT TTG 695  
 Ala Trp Arg Ala Leu Gly Gly Ile Arg Trp Gly Arg Arg Pro Arg Leu  
 20 25 30  
 ACC CCT GAC CTC CGG GCC CTG CTG ACG TCA GGA ACT TCT GAC CCC CGG 743  
 Thr Pro Asp Leu Arg Ala Leu Leu Thr Ser Gly Thr Ser Asp Pro Arg  
 35 40 45  
 GCC CGA GTG ACT TAT GGG ACC CCC AGT CTC TGG GCC CGG TTG TCT GTT 791  
 Ala Arg Val Thr Tyr Gly Thr Pro Ser Leu Trp Ala Arg Leu Ser Val  
 50 55 60  
 GGG GTC ACT GAA CCC CGA GCA TGC CTG ACG TCT GGG ACC CCG GGT CCC 839  
 Gly Val Thr Glu Pro Arg Ala Cys Leu Thr Ser Gly Thr Pro Gly Pro  
 65 70 75  
 CGG GCA CAA CTG ACT GCG GTG ACC CCA GAT ACC AGG ACC CGG GAG GCC 887  
 Arg Ala Gln Leu Thr Ala Val Thr Pro Asp Thr Arg Thr Arg Glu Ala  
 80 85 90 95  
 TCA GAG AAC TCT GGA ACC CGT TCG CGC GCG TGG CTG GCG GTG GCG CTG 935  
 Ser Glu Asn Ser Gly Thr Arg Ser Arg Ala Trp Leu Ala Val Ala Leu  
 100 105 110  
 GGC GCT GGG GGG GCA GTG CTG TTG TTG TGG GGC GGG GGT CCG GGT 983  
 Gly Ala Gly Gly Ala Val Leu Leu Leu Trp Gly Gly Gly Arg Gly  
 115 120 125

CCT CCG GCC GTC CTC GCC GCC GTC CCT AGC CCG CCG CCC GCT TCT CCC 1031  
 Pro Pro Ala Val Leu Ala Ala Val Pro Ser Pro Pro Pro Ala Ser Pro  
 130 135 140  
 CGG AGT CAG TAC AAC TTC ATC GCA GAT GTG GTG GAG AAG ACA GCA CCT 1079  
 Arg Ser Gln Tyr Asn Phe Ile Ala Asp Val Val Glu Lys Thr Ala Pro  
 145 150 155  
 GCC GTG GTC TAT ATC GAG ATC CTG GAC CGG CAC CCT TTC TTG GGC CGC 1127  
 Ala Val Val Tyr Ile Glu Ile Leu Asp Arg His Pro Phe Leu Gly Arg  
 160 165 170 175  
 GAG GTC CCT ATC TCG AAC GGC TCA GGA TTC GTG GTG GCT GCC GAT GGG 1175  
 Glu Val Pro Ile Ser Asn Gly Ser Gly Phe Val Val Ala Ala Asp Gly

65	180	185	190	66
CTC ATT GTC ACC AAC GCC CAT GTG GTG GCT GAT CGG CGC AGA GTC CGT				1223
Leu Ile Val Thr Asn Ala His Val Val Ala Asp Arg Arg Arg Val Arg				
195	200	205		
GTG AGA CTG CTA AGC GGC GAC ACG TAT GAG GCC GTG GTC ACA GCT GTG				1271
Val Arg Leu Leu Ser Gly Asp Thr Tyr Glu Ala Val Val Thr Ala Val				
210	215	220		
GAT CCC GTG GCA GAC ATC GCA ACG CTG AGG ATT CAG ACT AAG GAG CCT				1319
Asp Pro Val Ala Asp Ile Ala Thr Leu Arg Ile Gln Thr Lys Glu Pro				
225	230	235		
CTC CCC ACG CTG CCT CTG GGA CGC TCA GCT GAT GTC CGG CAA GGG GAG				1367
Leu Pro Thr Leu Pro Leu Gly Arg Ser Ala Asp Val Arg Gln Gly Glu				
240	245	250	255	
TTT GTT GTT GCC ATG GGA AGT CCC TTT GCA CTG CAG AAC ACG ATC ACA				1415
Phe Val Val Ala Met Gly Ser Pro Phe Ala Leu Gln Asn Thr Ile Thr				
260	265	270		
TCC GGC ATT GTT AGC TCT GCT CAG CGT CCA GCC AGA GAC CTG GGA CTC				1463
Ser Gly Ile Val Ser Ser Ala Gln Arg Pro Ala Arg Asp Leu Gly Leu				
275	280	285		
CCC CAA ACC AAT GTG GAA TAC ATT CAA ACT GAT GCA GCT ATT GAT TTT				1511
Pro Gln Thr Asn Val Glu Tyr Ile Gln Thr Asp Ala Ala Ile Asp Phe				
290	295	300		
GGA AAC TCT GGA GGT CCC CTG GTT AAC CTG GAT GGG GAG GTG ATT GGA				1559
Gly Asn Ser Gly Gly Pro Leu Val Asn Leu Asp Gly Glu Val Ile Gly				
305	310	315		
GTG AAC ACC ATG AAG GTC ACA GCT GGA ATC TCC TTT GCC ATC CCT TCT				1607
Val Asn Thr Met Lys Val Thr Ala Gly Ile Ser Phe Ala Ile Pro Ser				
320	325	330	335	
GAT CGT CTT CGA GAG TTT CTG CAT CGT GGG GAA AAG AAG AAT TCC TCC				1655
Asp Arg Leu Arg Glu Phe Leu His Arg Gly Glu Lys Lys Asn Ser Ser				
340	345	350		
TCC GGA ATC AGT GGG TCC CAG CGG CGC TAC ATT GGG GTG ATG ATG CTG				1703
Ser Gly Ile Ser Gly Ser Gln Arg Arg Tyr Ile Gly Val Met Met Leu				
355	360	365		
ACC CTG AGT CCC AGC ATC CTT GCT GAA CTA CAG CTT CGA GAA CCA AGC				1751
Thr Leu Ser Pro Ser Ile Leu Ala Glu Leu Gln Leu Arg Glu Pro Ser				
370	375	380		
TTT CCC GAT GTT CAG CAT GGT GTA CTC ATC CAT AAA GTC ATC CTG GGC				1799
Phe Pro Asp Val Gln His Gly Val Leu Ile His Lys Val Ile Leu Gly				
385	390	395		
TCC CCT GCA CAC CGG GCT GGT CTG CGG CCT GGT GAT GTG ATT TTG GCC				1847
Ser Pro Ala His Arg Ala Gly Leu Arg Pro Gly Asp Val Ile Leu Ala				
400	405	410	415	
ATT GGG GAG CAG ATG GTA CAA AAT GCT GAA GAT GTT TAT GAA GCT GTT				1895
Ile Gly Glu Gln Met Val Gln Asn Ala Glu Asp Val Tyr Glu Ala Val				
420	425	430		
CGA ACC CAA TCC CAG TTG GCA GTG CAG ATC CGG CGG GGA CGA GAA ACA				943
Arg Thr Gln Ser Gln Leu Ala Val Gln Ile Arg Arg Gly Arg Glu Thr				
435	440	445		
CTG ACC TTA TAT GTG ACC CCT GAG GTC ACA GAA TGAATAGATC ACCAAGAGTA				1996

67

68

Leu Thr Leu Tyr Val Thr Pro Glu Val Thr Glu

450

455

TGAGGCTCCT GCTCTGATTT CCTCCTTGCC TTTCTGGCTG AGGTTCTGAG GGCACCGAGA 2056

CAGAGGGTTA AATGAACCAG TGGGGGCAGG TCCCTCCAAC CACCAGCACT GACTCCTGGG 2116

CTCTGAAGAA TCACAGAAAC ACTTTTATA TAAATAAAA TTATACCTAG CAACATAAAA 2176

AAAAAAAAA A

2187

【0095】

配列番号: 25

Met Ala Ala Pro Arg Ala Gly Arg Gly Ala Gly Trp Ser Leu Arg Ala

1

5

10

15

Trp Arg Ala Leu Gly Gly Ile Arg Trp Gly Arg Arg Pro Arg Leu Thr

20

25

30

Pro Asp Leu Arg Ala Leu Leu Thr Ser Gly Thr Ser Asp Pro Arg Ala

35

40

45

Arg Val Thr Tyr Gly Thr Pro Ser Leu Trp Ala Arg Leu Ser Val Gly

50

55

60

Val Thr Glu Pro Arg Ala Cys Leu Thr Ser Gly Thr Pro Gly Pro Arg

65

70

75

80

Ala Gln Leu Thr Ala Val Thr Pro Asp Thr Arg Thr Arg Glu Ala Ser

85

90

95

Glu Asn Ser Gly Thr Arg Ser Arg Ala Trp Leu Ala Val Ala Leu Gly

100

105

110

Ala Gly Gly Ala Val Leu Leu Leu Leu Trp Gly Gly Gly Arg Gly Pro

115

120

125

Pro Ala Val Leu Ala Ala Val Pro Ser Pro Pro Pro Ala Ser Pro Arg

130

135

140

Ser Gln Tyr Asn Phe Ile Ala Asp Val Val Glu Lys Thr Ala Pro Ala

145

150

155

160

Val Val Tyr Ile Glu Ile Leu Asp Arg His Pro Phe Leu Gly Arg Glu

165

170

175

Val Pro Ile Ser Asn Gly Ser Gly Phe Val Val Ala Ala Asp Gly Leu

180

185

190

Ile Val Thr Asn Ala His Val Val Ala Asp Arg Arg Arg Val Arg Val

195

200

205

Arg Leu Leu Ser Gly Asp Thr Tyr Glu Ala Val Val Thr Ala Val Asp

210

215

220

Pro Val Ala Asp Ile Ala Thr Leu Arg Ile Gln Thr Lys Glu Pro Leu

225

230

235

240

Pro Thr Leu Pro Leu Gly Arg Ser Ala Asp Val Arg Gln Gly Glu Phe

245

250

255

Val Val Ala Met Gly Ser Pro Phe Ala Leu Gln Asn Thr Ile Thr Ser

260

265

270

Gly Ile Val Ser Ser Ala Gln Arg Pro Ala Arg Asp Leu Gly Leu Pro

275

280

285

Gln Thr Asn Val Glu Tyr Ile Gln Thr Asp Ala Ala Ile Asp Phe Gly

290

295

300

Asn Ser Gly Gly Pro Leu Val Asn Leu Asp Gly Glu Val Ile Gly Val

305

310

315

320

Asn Thr Met Lys Val Thr Ala Gly Ile Ser Phe Ala Ile Pro Ser Asp

325

330

335

69 70  
 Arg Leu Arg Glu Phe Leu His Arg Gly Glu Lys Lys Asn Ser Ser Ser  
 340 345 350  
 Gly Ile Ser Gly Ser Gln Arg Arg Tyr Ile Gly Val Met Met Leu Thr  
 355 360 365  
 Leu Ser Pro Ser Ile Leu Ala Glu Leu Gln Leu Arg Glu Pro Ser Phe  
 370 375 380  
 Pro Asp Val Gln His Gly Val Leu Ile His Lys Val Ile Leu Gly Ser  
 385 390 395 400  
 Pro Ala His Arg Ala Gly Leu Arg Pro Gly Asp Val Ile Leu Ala Ile  
 405 410 415  
 Gly Glu Gln Met Val Gln Asn Ala Glu Asp Val Tyr Glu Ala Val Arg  
 420 425 430  
 Thr Gln Ser Gln Leu Ala Val Gln Ile Arg Arg Gly Arg Glu Thr Leu  
 435 440 445  
 Thr Leu Tyr Val Thr Pro Glu Val Thr Glu  
 450 455

[0096]

配列番号: 26

CGTGGATCCC GAGAAAGAGG CGCAGGACGA GGAGGCAGAA CCCGACTGGC GCGTAGAGCA 60  
 GCAGCAGGAG CAGTAGGAAG CAGTCACCCG GAAGCCTGGG GGCGAGAGGC GAAGTGGTCA 120  
 GGCGCCGAAG GCCGAGAGCA CGCGGGGATC GGTCTCTTCC CGCCGGGTCT CTTACCGGTG 180  
 CGAGTCAAAG AGCCGCTCCG GCGCCGGCCC TGAGGGAAGC TCCATAACTG CTGCTTCAGG 240  
 AGCGCCCGGC CGTCGCCGCC GCGGCCATTT TCGCGCCCGG CCGCAGGGGC TCTTGGGAAG 300  
 GCGGAGTCTT TGGGCATCCG CCCGGGGTGA GGGGACCCGA AGTCCTGAGG CGCGCCGGAA 360  
 GGGCTAGCGG TCCCAGCATA CCCCGCGGCC CCTTGGGCCG TCTCACAAC TCGCTCCGGC 420  
 GGAGACCACA ATTCGCGGCA TTCGTGGGGC AGGGAGGAGT CGGCCTCCCG GAATCCTGGT 480  
 CCCGCGGTGC ACTTCTGAAG GACTTCAGGT ACCGGCGTGC CCCGCGTCCT ACTGTCCGCC 540  
 TGCTCGCGTC CTGGGTGCCG CCTCTGAGTA GGGCGGGCGA GGAGGCAGCC AAGGCGGAGC 600  
 TG ATG GCT GCG CCG AGG GCG GGG CGG GGT GCA GGC TGG AGC CTT CGG 647  
 Met Ala Ala Pro Arg Ala Gly Arg Gly Ala Gly Trp Ser Leu Arg  
 1 5 10 15  
 GCA TGG CGG GCT TTG GGG GGC ATT CGC TGG GGG AGG AGA CCC CGT TTG 695  
 Ala Trp Arg Ala Leu Gly Gly Ile Arg Trp Gly Arg Arg Pro Arg Leu  
 20 25 30  
 ACC CCT GAC CTC CGG GCC CTG CTG ACG TCA GGA ACT TCT GAC CCC CGG 743  
 Thr Pro Asp Leu Arg Ala Leu Leu Thr Ser Gly Thr Ser Asp Pro Arg  
 35 40 45  
 GCC CGA GTG ACT TAT GGG ACC CCC AGT CTC TGG GCC CGG TTG TCT GTT 791  
 Ala Arg Val Thr Tyr Gly Thr Pro Ser Leu Trp Ala Arg Leu Ser Val  
 50 55 60  
 GGG GTC ACT GAA CCC CGA GCA TGC CTG ACG TCT GGG ACC CCG GGT CCC 839  
 Gly Val Thr Glu Pro Arg Ala Cys Leu Thr Ser Gly Thr Pro Gly Pro  
 65 70 75  
 CGG GCA CAA CTG ACT GCG GTG ACC CCA GAT ACC AGG ACC CGG GAG GCC 887  
 Arg Ala Gln Leu Thr Ala Val Thr Pro Asp Thr Arg Thr Arg Glu Ala  
 80 85 90 95  
 TCA GAG AAC TCT GGA ACC CGT TCG CGC GCG TGG CTG GCG GTG GCG CTG 935  
 Ser Glu Asn Ser Gly Thr Arg Ser Arg Ala Trp Leu Ala Val Ala Leu  
 100 105 110  
 GGC GCT GGG GGG GCA GTG CTG TTG TTG TTG TGG GGC GGG GGT CGG GGT 983

71	72
Gly Ala Gly Gly Ala Val Leu Leu Leu Leu Trp Gly Gly Gly Arg Gly	
115	120 125
CCT CCG GCC GTC CTC GCC GCC GTC CCT AGC CCG CCG CCC GCT TCT CCC	1031
Pro Pro Ala Val Leu Ala Ala Val Pro Ser Pro Pro Pro Ala Ser Pro	
130	135 140
CGG AGT CAG TAC AAC TTC ATC GCA GAT GTG GTG GAG AAG ACA GCA CCT	1079
Arg Ser Gln Tyr Asn Phe Ile Ala Asp Val Val Glu Lys Thr Ala Pro	
145	150 155
GCC GTG GTC TAT ATC GAG ATC CTG GAC CGG CAC CCT TTC TTG GGC CGC	1127
Ala Val Val Tyr Ile Glu Ile Leu Asp Arg His Pro Phe Leu Gly Arg	
160	165 170 175
GAG GTC CCT ATC TCG AAC GGC TCA GGA TTC GTG GTG GCT GCC GAT GGG	1175
Glu Val Pro Ile Ser Asn Gly Ser Gly Phe Val Val Ala Ala Asp Gly	
180	185 190
CTC ATT GTC ACC AAC GCC CAT GTG GTG GCT GAT CGG CGC AGA GTC CGT	1223
Leu Ile Val Thr Asn Ala His Val Val Ala Asp Arg Arg Arg Val Arg	
195	200 205
GTG AGA CTG CTA AGC GGC GAC ACG TAT GAG GCC GTG GTC ACA GCT GTG	1271
Val Arg Leu Leu Ser Gly Asp Thr Tyr Glu Ala Val Val Thr Ala Val	
210	215 220
GAT CCC GTG GCA GAC ATC GCA ACG CTG AGG ATT CAG ACT AAG GAG CCT	1319
Asp Pro Val Ala Asp Ile Ala Thr Leu Arg Ile Gln Thr Lys Glu Pro	
225	230 235
CTC CCC ACG CTG CCT CTG GGA CGC TCA GCT GAT GTC CGG CAA GGG GAG	1367
Leu Pro Thr Leu Pro Leu Gly Arg Ser Ala Asp Val Arg Gln Gly Glu	
240	245 250 255
TTT GTT GTT GCC ATG GGA AGT CCC TTT GCA CTG CAG AAC ACG ATC ACA	1415
Phe Val Val Ala Met Gly Ser Pro Phe Ala Leu Gln Asn Thr Ile Thr	
260	265 270
TCC GGC ATT GTT AGC TCT GCT CAG CGT CCA GCC AGA GAC CTG GGA CTC	1463
Ser Gly Ile Val Ser Ser Ala Gln Arg Pro Ala Arg Asp Leu Gly Leu	
275	280 285
CCC CAA ACC AAT GTG GAA TAC ATT CAA ACT GAT GCA GCT ATT GAT TTT	1511
Pro Gln Thr Asn Val Glu Tyr Ile Gln Thr Asp Ala Ala Ile Asp Phe	
290	295 300
GGA AAC TCT GGA GGT CCC CTG GTT AAC CTG GTG AGT GAG ACA TCC TTC	1559
Gly Asn Ser Gly Gly Pro Leu Val Asn Leu Val Ser Glu Thr Ser Phe	
305	310 315
CTT CCA AGA ATC CCT GCC CCA GGT CAG TGT GGG AAG GGT AGG TTT CCC	1607
Leu Pro Arg Ile Pro Ala Pro Gly Gln Cys Gly Lys Gly Arg Phe Pro	
320	325 330 335
CTA ATT CAA GGA TGT TTG GTC AAG TTT CTG AGC AGT TCT TTG TTG GCT	1655
Leu Ile Gln Gly Cys Leu Val Lys Phe Leu Ser Ser Ser Leu Leu Ala	
340	345 350
ATC TCT CAA TAT CCA ACC AGA TCT CCC CAA CAC TTG CTG GTA CTT TTG	1703
Ile Ser Gln Tyr Pro Thr Arg Ser Pro Gln His Leu Leu Val Leu Leu	
355	360 365
TTC GGG TGC CCC CAT CCC CTA CTA TTT GTT TAGGCTAGGG AACTGGGGGC TGTA	1757
Phe Gly Cys Pro His Pro Leu Leu Phe Val	
370	375

73

74

TCCCTGCAGG ATGGGGAGGT GATTGGAGTG AACACCATGA AGGTCACAGC TGGAAATCTCC 1817  
 TTTGCCATCC CTTCTGATCG TCTTCGAGAG TTTCTGCATC GTGGGGAAAA GAAGAATTCC 1877  
 TCCTCCGGAA TCAGTGGGTC CCAGCGGCGC TACATTGGGG TGATGATGCT GACCCTGAGT 1937  
 CCCAGCATCC TTGTGAACT ACAGCTTCGA GAACCAAGCT TTCCCGATGT TCAGCATGGT 1997  
 GTACTCATCC ATAAAGTCAT CCTGGGCTCC CCTGCACACC GGGCTGGTCT GCGGCCTGGT 2057  
 GATGTGATTT TGGCCATTGG GGAGCAGATG GTACAAAATG CTGAAGATGT TTATGAAGCT 2117  
 GTTCGAACCC AATCCCAGTT GGCAGTGCAG ATCCGGCGGG GACGAGAAAC ACTGACCTTA 2177  
 TATGTGACCC CTGAGGTCAC AGAATGAATA GATCACCAAG AGTATGAGGC TCCTGCTCTG 2237  
 ATTTCTCTCT TGCCTTTCTG GCTGAGGTTC TGAGGGCACC GAGACAGAGG GTTAAATGAA 2297  
 CCAGTGGGGG CAGGTCCCTC CAACCACCAG CACTGACTCC TGGGCTCTGA AGAATCACAG 2357  
 AAACACTTTT TATATAAAAT AAAATTATAC CTAGCAACAT ATTATAGTAA AAAATGAGGT 2417  
 GGGAGGGCTG GATCTTTTCC CCCACAAAA GGCTAGAGGT AAAGCTGTAT CCCCTAAAC 2477  
 TTAGGGGAGA TACTGGAGCT GACCATCTG ACCTCCTATT AAAGAAAATG AGCTGCTGAA 2537  
 AAAAAAAAAA AAAA 2551

【0097】

配列番号：27

Met Ala Ala Pro Arg Ala Gly Arg Gly Ala Gly Trp Ser Leu Arg Ala  
 1 5 10 15  
 Trp Arg Ala Leu Gly Gly Ile Arg Trp Gly Arg Arg Pro Arg Leu Thr  
 20 25 30  
 Pro Asp Leu Arg Ala Leu Leu Thr Ser Gly Thr Ser Asp Pro Arg Ala  
 35 40 45  
 Arg Val Thr Tyr Gly Thr Pro Ser Leu Trp Ala Arg Leu Ser Val Gly  
 50 55 60  
 Val Thr Glu Pro Arg Ala Cys Leu Thr Ser Gly Thr Pro Gly Pro Arg  
 65 70 75 80  
 Ala Gln Leu Thr Ala Val Thr Pro Asp Thr Arg Thr Arg Glu Ala Ser  
 85 90 95  
 Glu Asn Ser Gly Thr Arg Ser Arg Ala Trp Leu Ala Val Ala Leu Gly  
 100 105 110  
 Ala Gly Gly Ala Val Leu Leu Leu Leu Trp Gly Gly Gly Arg Gly Pro  
 115 120 125  
 Pro Ala Val Leu Ala Ala Val Pro Ser Pro Pro Pro Ala Ser Pro Arg  
 130 135 140  
 Ser Gln Tyr Asn Phe Ile Ala Asp Val Val Glu Lys Thr Ala Pro Ala  
 145 150 155 160  
 Val Val Tyr Ile Glu Ile Leu Asp Arg His Pro Phe Leu Gly Arg Glu  
 165 170 175  
 Val Pro Ile Ser Asn Gly Ser Gly Phe Val Val Ala Ala Asp Gly Leu  
 180 185 190  
 Ile Val Thr Asn Ala His Val Val Ala Asp Arg Arg Arg Val Arg Val  
 195 200 205  
 Arg Leu Leu Ser Gly Asp Thr Tyr Glu Ala Val Val Thr Ala Val Asp  
 210 215 220  
 Pro Val Ala Asp Ile Ala Thr Leu Arg Ile Gln Thr Lys Glu Pro Leu  
 225 230 235 240  
 Pro Thr Leu Pro Leu Gly Arg Ser Ala Asp Val Arg Gln Gly Glu Phe  
 245 250 255  
 Val Val Ala Met Gly Ser Pro Phe Ala Leu Gln Asn Thr Ile Thr Ser  
 260 265 270



75 76  
 Gly Ile Val Ser Ser Ala Gln Arg Pro Ala Arg Asp Leu Gly Leu Pro  
 275 280 285  
 Gln Thr Asn Val Glu Tyr Ile Gln Thr Asp Ala Ala Ile Asp Phe Gly  
 290 295 300  
 Asn Ser Gly Gly Pro Leu Val Asn Leu Val Ser Glu Thr Ser Phe Leu  
 305 310 315 320  
 Pro Arg Ile Pro Ala Pro Gly Gln Cys Gly Lys Gly Arg Phe Pro Leu  
 325 330 335  
 Ile Gln Gly Cys Leu Val Lys Phe Leu Ser Ser Ser Leu Leu Ala Ile  
 340 345 350  
 Ser Gln Tyr Pro Thr Arg Ser Pro Gln His Leu Leu Val Leu Leu Phe  
 355 360 365  
 Gly Cys Pro His Pro Leu Leu Phe Val  
 370 375

【0098】

配列番号：28

CGTGGATCCC GAGAAAGAGG CGCAGGACGA GGAGGCAGAA CCCGACTGGC GCGTAGAGCA 60  
 GCAGCACGAG CAGTAGGAAG CAGTCACCCG GAAGCCTGGG GCGGAGAGGC GAAGTGGTCA 120  
 GGCGCCGAAG GCCGAGAGCA CGCGGGGATC GGTCTCTTCC CGCCGGGTCT CTTACCGGTG 180  
 CGAGTCAAAG AGCCGCTCCG GCCCGGGCCC TGAGGGAAGC TCCATAACTG CTGCTTCAGG 240  
 AGCGCCCGGC CGTCGCCGCC GCCGCCATTT TCGCGCCCGG CCGCAGGGGC TCTTGGGAAG 300  
 GCGGAGTCTT TGGGCATCCG CCCGGGGTGA GGGGACCCGA AGTCTGAGG CGCGCCGGAA 360  
 GGGCTAGCGG TCCCAGCATA CCCCGCGGCC CCTTGGGCCG TCTCACAAC TCGGTCCGGC 420  
 GGAGACCACA ATTCCCGGCA TTCGTGGGGC AGGAGGAGT CGGCCTCCCG GAATCCTGGT 480  
 CCCGGCGTGC ACTTCTGAAG GACTTCAGGT ACCGGCGTGC CCCGCGTCCT ACTGTCCGCC 540  
 TGCTCGCGTC CTGGGTGCCG CCTCTGAGTA GGGCGGGCGA GGAGGCAGCC AAGGCGGAGC 600  
 TG ATG GCT GCG CCG AGG GCG GGG CGG GGT GCA GGC TGG AGC CTT CGG 647  
 Met Ala Ala Pro Arg Ala Gly Arg Gly Ala Gly Trp Ser Leu Arg  
 1 5 10 15  
 GCA TGG CGG GCT TTG GGG GGC ATT CGC TGG GGG AGG AGA CCC CGT TTG 695  
 Ala Trp Arg Ala Leu Gly Gly Ile Arg Trp Gly Arg Arg Pro Arg Leu  
 20 25 30  
 ACC CCT GAC CTC CGG GCC CTG CTG ACG TCA GGA ACT TCT GAC CCC CGG 743  
 Thr Pro Asp Leu Arg Ala Leu Leu Thr Ser Gly Thr Ser Asp Pro Arg  
 35 40 45  
 GCC CGA GTG ACT TAT GGG ACC CCC AGT CTC TGG GCC CGG TTG TCT GTT 791  
 Ala Arg Val Thr Tyr Gly Thr Pro Ser Leu Trp Ala Arg Leu Ser Val  
 50 55 60  
 GGG GTC ACT GAA CCC CGA GCA TGC CTG ACG TCT GGG ACC CCG GGT CCC 839  
 Gly Val Thr Glu Pro Arg Ala Cys Leu Thr Ser Gly Thr Pro Gly Pro  
 65 70 75  
 CGG GCA CAA CTG ACT GCG GTG ACC CCA GAT ACC AGG ACC CGG GAG GCC 887  
 Arg Ala Gln Leu Thr Ala Val Thr Pro Asp Thr Arg Thr Arg Glu Ala  
 80 85 90 95  
 TCA GAG AAC TCT GGA ACC CGT TCG CGC GCG TGG CTG GCG GTG GCG CTG 935  
 Ser Glu Asn Ser Gly Thr Arg Ser Arg Ala Trp Leu Ala Val Ala Leu  
 100 105 110  
 GGC GCT GGG GGG GCA GTG CTG TTG TTG TTG TGG GGC GGG GGT CGG GGT 983  
 Gly Ala Gly Gly Ala Val Leu Leu Leu Leu Trp Gly Gly Gly Arg Gly  
 115 120 125

77	78	
CCT CCG GCC GTC CTC GCC GCC GTC CCT AGC CCG CCG CCC GCT TCT CCC	1031	
Pro Pro Ala Val Leu Ala Ala Val Pro Ser Pro Pro Pro Ala Ser Pro		
130 135 140		
CGG AGT CAG TAC AAC TTC ATC GCA GAT GTG GTG GAG AAG ACA GCA CCT	1079	
Arg Ser Gln Tyr Asn Phe Ile Ala Asp Val Val Glu Lys Thr Ala Pro		
145 150 155		
GCC GTG GTC TAT ATC GAG ATC CTG GAC CGG CAC CCT TTC TTG GGC CGC	1127	
Ala Val Val Tyr Ile Glu Ile Leu Asp Arg His Pro Phe Leu Gly Arg		
160 165 170 175		
GAG GTC CCT ATC TCG AAC GGC TCA GGA TTC GTG GTG GCT GCC GAT GGG	1175	
Glu Val Pro Ile Ser Asn Gly Ser Gly Phe Val Val Ala Ala Asp Gly		
180 185 190		
CTC ATT GTC ACC AAC GCC CAT GTG GTG GCT GAT CGG CGC AGA GTC CGT	1223	
Leu Ile Val Thr Asn Ala His Val Val Ala Asp Arg Arg Arg Val Arg		
195 200 205		
GTG AGA CTG CTA AGC GGC GAC ACG TAT GAG GCC GTG GTC ACA GCT GTG	1271	
Val Arg Leu Leu Ser Gly Asp Thr Tyr Glu Ala Val Val Thr Ala Val		
210 215 220		
GAT CCC GTG GCA GAC ATC GCA ACG CTG AGG ATT CAG ACT AAG GAG CCT	1319	
Asp Pro Val Ala Asp Ile Ala Thr Leu Arg Ile Gln Thr Lys Glu Pro		
225 230 235		
CTC CCC ACG CTG CCT CTG GGA CGC TCA GCT GAT GTC CGG CAA GGG GAG	1367	
Leu Pro Thr Leu Pro Leu Gly Arg Ser Ala Asp Val Arg Gln Gly Glu		
240 245 250 255		
TTT GTT GTT GCC ATG GGA AGT CCC TTT GCA CTG CAG AAC ACG ATC ACA	1415	
Phe Val Val Ala Met Gly Ser Pro Phe Ala Leu Gln Asn Thr Ile Thr		
260 265 270		
TCC GGC ATT GTT AGC TCT GCT CAG CGT CCA GCC AGA GAC CTG GGA CTC	1463	
Ser Gly Ile Val Ser Ser Ala Gln Arg Pro Ala Arg Asp Leu Gly Leu		
275 280 285		
CCC CAA ACC AAT GTG GAA TAC ATT CAA ACT GAT GCA GCT ATT GAT TTT	1511	
Pro Gln Thr Asn Val Glu Tyr Ile Gln Thr Asp Ala Ala Ile Asp Phe		
290 295 300		
GGA AAC TCT GGA GGT CCC CTG GTT AAC CTG GCT AGG GAA CTG GGG GCT	1559	
Gly Asn Ser Gly Gly Pro Leu Val Asn Leu Ala Arg Glu Leu Gly Ala		
305 310 315		
GTA TCC CTG CAG GAT GGG GAG GTG ATT GGA GTG AAC ACC ATG AAG GTC	1607	
Val Ser Leu Gln Asp Gly Glu Val Ile Gly Val Asn Thr Met Lys Val		
320 325 330 335		
ACA GCT GGA ATC TCC TTT GCC ATC CCT TCT GAT CGT CTT CGA GAG TTT	1655	
Thr Ala Gly Ile Ser Phe Ala Ile Pro Ser Asp Arg Leu Arg Glu Phe		
340 345 350		
CTG CAT CGT GGG GAA AAG AAG AAT TCC TCC TCC GGA ATC AGT GGG TCC	1703	
Leu His Arg Gly Glu Lys Lys Asn Ser Ser Ser Gly Ile Ser Gly Ser		
355 360 365		
CAG CGG CGC TAC ATT GGG GTG ATG ATG CTG ACC CTG AGT CCC AGG GCT	1751	
Gln Arg Arg Tyr Ile Gly Val Met Met Leu Thr Leu Ser Pro Arg Ala		
370 375 380		
GGT CTG CGG CCT GGT GAT GTG ATT TTG GCC ATT GGG GAG CAG ATG GTA	1799	
Gly Leu Arg Pro Gly Asp Val Ile Leu Ala Ile Gly Glu Gln Met Val		

79 80  
 385 390 395  
 CAA AAT GCT GAA GAT GTT TAT GAA GCT GTT CGA ACC CAA TCC CAG TTG 1847  
 Gln Asn Ala Glu Asp Val Tyr Glu Ala Val Arg Thr Gln Ser Gln Leu  
 400 405 410 415  
 GCA GTG CAG ATC CGG CGG GGA CGA GAA ACA CTG ACC TTA TAT GTG ACC 1895  
 Ala Val Gln Ile Arg Arg Gly Arg Glu Thr Leu Thr Leu Tyr Val Thr  
 420 425 430  
 CCT GAG GTC ACA GAA TGAATAGATC ACCAAGAGTA TGAGGCTCCT GCTCTGATTT CC 1952  
 Pro Glu Val Thr Glu  
 435  
 TCCTTGCCCTT TCTGGCTGAG GTTCTGAGGG CACCGAGACA GAGGGTTAAA TGAACCACTG 2012  
 GGGGCAGGTC CCTCCAACCA CCAGCACTGA CTCCTGGGCT CTGAAGAATC ACAGAAACAC 2072  
 TTTTATATA AAATAAAATT ATACCTAGCA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 2132  
 AAAAAAAAAA AA 2144

【0099】

配列番号：29

Met Ala Ala Pro Arg Ala Gly Arg Gly Ala Gly Trp Ser Leu Arg Ala  
 1 5 10 15  
 Trp Arg Ala Leu Gly Gly Ile Arg Trp Gly Arg Arg Pro Arg Leu Thr  
 20 25 30  
 Pro Asp Leu Arg Ala Leu Leu Thr Ser Gly Thr Ser Asp Pro Arg Ala  
 35 40 45  
 Arg Val Thr Tyr Gly Thr Pro Ser Leu Trp Ala Arg Leu Ser Val Gly  
 50 55 60  
 Val Thr Glu Pro Arg Ala Cys Leu Thr Ser Gly Thr Pro Gly Pro Arg  
 65 70 75 80  
 Ala Gln Leu Thr Ala Val Thr Pro Asp Thr Arg Thr Arg Glu Ala Ser  
 85 90 95  
 Glu Asn Ser Gly Thr Arg Ser Arg Ala Trp Leu Ala Val Ala Leu Gly  
 100 105 110  
 Ala Gly Gly Ala Val Leu Leu Leu Leu Trp Gly Gly Gly Arg Gly Pro  
 115 120 125  
 Pro Ala Val Leu Ala Ala Val Pro Ser Pro Pro Pro Ala Ser Pro Arg  
 130 135 140  
 Ser Gln Tyr Asn Phe Ile Ala Asp Val Val Glu Lys Thr Ala Pro Ala  
 145 150 155 160  
 Val Val Tyr Ile Glu Ile Leu Asp Arg His Pro Phe Leu Gly Arg Glu  
 165 170 175  
 Val Pro Ile Ser Asn Gly Ser Gly Phe Val Val Ala Ala Asp Gly Leu  
 180 185 190  
 Ile Val Thr Asn Ala His Val Val Ala Asp Arg Arg Arg Val Arg Val  
 195 200 205  
 Arg Leu Leu Ser Gly Asp Thr Tyr Glu Ala Val Val Thr Ala Val Asp  
 210 215 220  
 Pro Val Ala Asp Ile Ala Thr Leu Arg Ile Gln Thr Lys Glu Pro Leu  
 225 230 235 240  
 Pro Thr Leu Pro Leu Gly Arg Ser Ala Asp Val Arg Gln Gly Glu Phe  
 245 250 255  
 Val Val Ala Met Gly Ser Pro Phe Ala Leu Gln Asn Thr Ile Thr Ser  
 260 265 270

81 82  
 Gly Ile Val Ser Ser Ala Gln Arg Pro Ala Arg Asp Leu Gly Leu Pro  
 275 280 285  
 Gln Thr Asn Val Glu Tyr Ile Gln Thr Asp Ala Ala Ile Asp Phe Gly  
 290 295 300  
 Asn Ser Gly Gly Pro Leu Val Asn Leu Ala Arg Glu Leu Gly Ala Val  
 305 310 315 320  
 Ser Leu Gln Asp Gly Glu Val Ile Gly Val Asn Thr Met Lys Val Thr  
 325 330 335  
 Ala Gly Ile Ser Phe Ala Ile Pro Ser Asp Arg Leu Arg Glu Phe Leu  
 340 345 350  
 His Arg Gly Glu Lys Lys Asn Ser Ser Ser Gly Ile Ser Gly Ser Gln  
 355 360 365  
 Arg Arg Tyr Ile Gly Val Met Met Leu Thr Leu Ser Pro Arg Ala Gly  
 370 375 380  
 Leu Arg Pro Gly Asp Val Ile Leu Ala Ile Gly Glu Gln Met Val Gln  
 385 390 395 400  
 Asn Ala Glu Asp Val Tyr Glu Ala Val Arg Thr Gln Ser Gln Leu Ala  
 405 410 415  
 Val Gln Ile Arg Arg Gly Arg Glu Thr Leu Thr Leu Tyr Val Thr Pro  
 420 425 430  
 Glu Val Thr Glu  
 435

【0100】配列番号：30

\*アミノ酸24=Arg/Cys アミノ酸278=Ala

特徴：672および1435で多型性変種

\* a/V a l

CGTGGATCCC GAGAAAGAGG CGCAGGACGA GGAGGCAGAA CCCGACTGGC GCGTAGAGCA 60  
 GCAGCACGAG CAGTAGGAAG CAGTCACCCG GAAGCCTGGG GGCAGAGGC GAAGTGGTCA 120  
 GCGGCCGAAG GCCGAGAGCA CGCGGGGATC GGTCTCTTCC CGCGGGGTCT CTTACCGGTG 180  
 CGAGTCAAAG AGCCGCTCCG GCCCGGCCCG TGAGGGAAGC TCCATACTG CTGCTTCAGG 240  
 AGCGCCCGGC CGTCGCCGCC GCCGCCATTT TCGCGCCCGG CCGCAGGGGC TCTTGGGAAG 300  
 GCGGAGTCTT TGGGCATCCG CCCGGGGTGA GGGGACCCGA AGTCCTGAGG CGCGCCGGAA 360  
 GGGCTAGCGG TCCCAGCATA CCCCGCGGCC CCTTGGGCCG TCTCACAACCT CGCGTCCGGC 420  
 GGAGACCACA ATTCCCGGCA TTCGTGGGGC AGGAGGAGT CGGCCTCCCG GAATCCTGGT 480  
 CCCGGCGTGC ACTTCTGAAG GACTTCAGGT ACCCGCGTGC CCCGCGTCCT ACTGTCCGCC 540  
 TGCTCGCGTC CTGGGTGCCG CCTCTGAGTA GGGCGGGCGA GGAGGCAGCC AAGGCGGAGC 600  
 TG ATG GCT GCG CCG AGG GCG GGG CGG GGT GCA GGC TGG AGC CTT CGG 647  
 Met Ala Ala Pro Arg Ala Gly Arg Gly Ala Gly Trp Ser Leu Arg  
 1 5 10 15  
 GCA TGG CGG GCT TTG GGG GGC ATT YGC TGG GGG AGG AGA CCC CGT TTG 695  
 Ala Trp Arg Ala Leu Gly Gly Ile Xaa Trp Gly Arg Arg Pro Arg Leu  
 20 25 30  
 ACC CCT GAC CTC CGG GCC CTG CTG ACG TCA GGA ACT TCT GAC CCC CGG 743  
 Thr Pro Asp Leu Arg Ala Leu Leu Thr Ser Gly Thr Ser Asp Pro Arg  
 35 40 45  
 GCC CGA GTG ACT TAT GGG ACC CCC AGT CTC TGG GCC CGG TTG TCT GTT 791  
 Ala Arg Val Thr Tyr Gly Thr Pro Ser Leu Trp Ala Arg Leu Ser Val  
 50 55 60  
 GGG GTC ACT GAA CCC CGA GCA TGC CTG ACG TCT GGG ACC CCG GGT CCC 839  
 Gly Val Thr Glu Pro Arg Ala Cys Leu Thr Ser Gly Thr Pro Gly Pro  
 65 70 75  
 CGG GCA CAA CTG ACT GCG GTG ACC CCA GAT ACC AGG ACC CGG GAG GCC 887

83	84
Arg Ala Gln Leu Thr Ala Val Thr Pro Asp Thr Arg Thr Arg Glu Ala	
80	85 90 95
TCA GAG AAC TCT GGA ACC CGT TCG CGC GCG TGG CTG GCG GTG GCG CTG	935
Ser Glu Asn Ser Gly Thr Arg Ser Arg Ala Trp Leu Ala Val Ala Leu	
100	105 110
GGC GCT GGG GGG GCA GTG CTG TTG TTG TTG TGG GGC GGG GGT CGG GGT	983
Gly Ala Gly Gly Ala Val Leu Leu Leu Leu Trp Gly Gly Gly Arg Gly	
115	120 125
CCT CCG GCC GTC CTC GCC GCC GTC CCT AGC CCG CCG CCC GCT TCT CCC	1031
Pro Pro Ala Val Leu Ala Ala Val Pro Ser Pro Pro Pro Ala Ser Pro	
130	135 140
CGG AGT CAG TAC AAC TTC ATC GCA GAT GTG GTG GAG AAG ACA GCA CCT	1079
Arg Ser Gln Tyr Asn Phe Ile Ala Asp Val Val Glu Lys Thr Ala Pro	
145	150 155
GCC GTG GTC TAT ATC GAG ATC CTG GAC CGG CAC CCT TTC TTG GGC CGC	1127
Ala Val Val Tyr Ile Glu Ile Leu Asp Arg His Pro Phe Leu Gly Arg	
160	165 170 175
GAG GTC CCT ATC TCG AAC GGC TCA GGA TTC GTG GTG GCT GCC GAT GGG	1175
Glu Val Pro Ile Ser Asn Gly Ser Gly Phe Val Val Ala Ala Asp Gly	
180	185 190
CTC ATT GTC ACC AAC GCC CAT GTG GTG GCT GAT CGG CGC AGA GTC CGT	1223
Leu Ile Val Thr Asn Ala His Val Val Ala Asp Arg Arg Arg Val Arg	
195	200 205
GTG AGA CTG CTA AGC GGC GAC ACG TAT GAG GCC GTG GTC ACA GCT GTG	1271
Val Arg Leu Leu Ser Gly Asp Thr Tyr Glu Ala Val Val Thr Ala Val	
210	215 220
GAT CCC GTG GCA GAC ATC GCA ACG CTG AGG ATT CAG ACT AAG GAG CCT	1319
Asp Pro Val Ala Asp Ile Ala Thr Leu Arg Ile Gln Thr Lys Glu Pro	
225	230 235
CTC CCC ACG CTG CCT CTG GGA CGC TCA GCT GAT GTC CGG CAA GGG GAG	1367
Leu Pro Thr Leu Pro Leu Gly Arg Ser Ala Asp Val Arg Gln Gly Glu	
240	245 250 255
TTT GTT GTT GCC ATG GGA AGT CCC TTT GCA CTG CAG AAC ACG ATC ACA	1415
Phe Val Val Ala Met Gly Ser Pro Phe Ala Leu Gln Asn Thr Ile Thr	
260	265 270
TCC GGC ATT GTT AGC TCT YCT CAG CGT CCA GCC AGA GAC CTG GGA CTC	1463
Ser Gly Ile Val Ser Ser Xaa Gln Arg Pro Ala Arg Asp Leu Gly Leu	
275	280 285
CCC CAA ACC AAT GTG CAA TAC ATT CAA ACT GAT GCA GCT ATT GAT TTT	1511
Pro Gln Thr Asn Val Glu Tyr Ile Gln Thr Asp Ala Ala Ile Asp Phe	
290	295 300
GGA AAC TCT GGA GGT CCC CTG GTT AAC CTG GAT GGG CAG GTG ATT GGA	1559
Gly Asn Ser Gly Gly Pro Leu Val Asn Leu Asp Gly Glu Val Ile Gly	
305	310 315
GTG AAC ACC ATG AAG GTC ACA GCT GGA ATC TCC TTT GCC ATC CCT TCT	1607
Val Asn Thr Met Lys Val Thr Ala Gly Ile Ser Phe Ala Ile Pro Ser	
320	325 330 335
GAT CGT CTT CGA GAG TTT CTG CAT CGT GGG GAA AAG AAG AAT TCC TCC	1655
Asp Arg Leu Arg Glu Phe Leu His Arg Gly Glu Lys Lys Asn Ser Ser	
340	345 350

85	86	
TCC GGA ATC AGT GGG TCC CAG CGG CGC TAC ATT GGG GTG ATG ATG CTG	1703	
Ser Gly Ile Ser Gly Ser Gln Arg Arg Tyr Ile Gly Val Met Met Leu		
355	360	365
ACC CTG AGT CCC AGC ATC CTT GCT GAA CTA CAG CTT CGA GAA CCA AGC	1751	
Thr Leu Ser Pro Ser Ile Leu Ala Glu Leu Gln Leu Arg Glu Pro Ser		
370	375	380
TTT CCC GAT GTT CAG CAT GGT GTA CTC ATC CAT AAA GTC ATC CTG GGC	1799	
Phe Pro Asp Val Gln His Gly Val Leu Ile His Lys Val Ile Leu Gly		
385	390	395
TCC CCT GCA CAC CGG GCT GGT CTG CGG CCT GGT GAT GTG ATT TTG GCC	1847	
Ser Pro Ala His Arg Ala Gly Leu Arg Pro Gly Asp Val Ile Leu Ala		
400	405	410
ATT GGG GAG CAG ATG GTA CAA AAT GCT GAA GAT GTT TAT GAA GCT GTT	1895	
Ile Gly Glu Gln Met Val Gln Asn Ala Glu Asp Val Tyr Glu Ala Val		
420	425	430
CGA ACC CAA TCC CAG TTG GCA GTG CAG ATC CGG CGG GGA CGA GAA ACA	1943	
Arg Thr Gln Ser Gln Leu Ala Val Gln Ile Arg Arg Gly Arg Glu Thr		
435	440	445
CTG ACC TTA TAT GTG ACC CCT GAG GTC ACA GAA TGAATAGATC ACCAAGAGTA	1996	
Leu Thr Leu Tyr Val Thr Pro Glu Val Thr Glu		
450	455	
TGAGGCTCCT GCTCTGATTT CCTCCTTGCC TTTCTGGCTG AGGTTCTGAG GGCACCGAGA	2056	
CAGAGGGTTA AATGAACCAG TGGGGGCAGG TCCCTCCAAC CACCAGCACT GACTCCTGGG	2116	
CTCTGAAGAA TCACAGAAAC ACTTTTATA TAAATAAAAA TTATACCTAG CAACATAAAA	2176	
AAAAAAAAA A	2187	

【0101】配列番号: 31

\* 2 4 X a a = A r g または C y s

特徴

\* 2 7 8 X a a = A l a または V a l

Met Ala Ala Pro Arg Ala Gly Arg Gly Ala Gly Trp Ser Leu Arg Ala			
1	5	10	15
Trp Arg Ala Leu Gly Gly Ile Xaa Trp Gly Arg Arg Pro Arg Leu Thr			
20	25	30	
Pro Asp Leu Arg Ala Leu Leu Thr Ser Gly Thr Ser Asp Pro Arg Ala			
35	40	45	
Arg Val Thr Tyr Gly Thr Pro Ser Leu Trp Ala Arg Leu Ser Val Gly			
50	55	60	
Val Thr Glu Pro Arg Ala Cys Leu Thr Ser Gly Thr Pro Gly Pro Arg			
65	70	75	80
Ala Gln Leu Thr Ala Val Thr Pro Asp Thr Arg Thr Arg Glu Ala Ser			
85	90	95	
Glu Asn Ser Gly Thr Arg Ser Arg Ala Trp Leu Ala Val Ala Leu Gly			
100	105	110	
Ala Gly Gly Ala Val Leu Leu Leu Trp Gly Gly Gly Arg Gly Pro			
115	120	125	
Pro Ala Val Leu Ala Ala Val Pro Ser Pro Pro Pro Ala Ser Pro Arg			
130	135	140	
Ser Gln Tyr Asn Phe Ile Ala Asp Val Val Glu Lys Thr Ala Pro Ala			
145	150	155	160
Val Val Tyr Ile Glu Ile Leu Asp Arg His Pro Phe Leu Gly Arg Glu			
165	170	175	
Val Pro Ile Ser Asn Gly Ser Gly Phe Val Val Ala Ala Asp Gly Leu			

87 88

180 185 190

Ile Val Thr Asn Ala His Val Val Ala Asp Arg Arg Arg Val Arg Val

195 200 205

Arg Leu Leu Ser Gly Asp Thr Tyr Glu Ala Val Val Thr Ala Val Asp

210 215 220

Pro Val Ala Asp Ile Ala Thr Leu Arg Ile Gln Thr Lys Glu Pro Leu

225 230 235 240

Pro Thr Leu Pro Leu Gly Arg Ser Ala Asp Val Arg Gln Gly Glu Phe

245 250 255

Val Val Ala Met Gly Ser Pro Phe Ala Leu Gln Asn Thr Ile Thr Ser

260 265 270

Gly Ile Val Ser Ser Xaa Gln Arg Pro Ala Arg Asp Leu Gly Leu Pro

275 280 285

Gln Thr Asn Val Glu Tyr Ile Gln Thr Asp Ala Ala Ile Asp Phe Gly

290 295 300

Asn Ser Gly Gly Pro Leu Val Asn Leu Asp Gly Glu Val Ile Gly Val

305 310 315 320

Asn Thr Met Lys Val Thr Ala Gly Ile Ser Phe Ala Ile Pro Ser Asp

325 330 335

Arg Leu Arg Glu Phe Leu His Arg Gly Glu Lys Lys Asn Ser Ser Ser

340 345 350

Gly Ile Ser Gly Ser Gln Arg Arg Tyr Ile Gly Val Met Met Leu Thr

355 360 365

Leu Ser Pro Ser Ile Leu Ala Glu Leu Gln Leu Arg Glu Pro Ser Phe

370 375 380

Pro Asp Val Gln His Gly Val Leu Ile His Lys Val Ile Leu Gly Ser

385 390 395 400

Pro Ala His Arg Ala Gly Leu Arg Pro Gly Asp Val Ile Leu Ala Ile

405 410 415

Gly Glu Gln Met Val Gln Asn Ala Glu Asp Val Tyr Glu Ala Val Arg

420 425 430

Thr Gln Ser Gln Leu Ala Val Gln Ile Arg Arg Gly Arg Glu Thr Leu

435 440 445

Thr Leu Tyr Val Thr Pro Glu Val Thr Glu

450 455

【0102】配列番号: 32

CATCCGGCAT TGTTAGCTCT GC 22

【0103】配列番号: 33

CATCCGGCAT-TGTTAGCTCT-GT 22

【0104】配列番号: 34

CAATAGCTGC ATCAGTTTGA ATG 23

【0105】配列番号: 35

TGGCGGGCTT TGGGGGGCAT TC 22

【0106】配列番号: 36

TGGCGGGCTT TGGGGGGCAT TT 22

【0107】配列番号: 37

GACGTCAGCA GGGCCCGGAG GTC 23

【0108】配列番号: 38

GATACCCAG CAGAAGCTGG 20

【0109】配列番号: 39

GATACCCAG CAGAAGCTGT 20

【0110】配列番号: 40

GCTGACATCA TTGGCGGAGA C 21

【図面の簡単な説明】

40 【図1】 イー・コリ (E.coli) htrAでPSP-1  
のアミノ酸配列を整列化したものである。【図2】 PSP1単離体PSP1-1、PSP1-  
2、PSP1-3およびPSP1-4の複数のcDNA  
配列を整列化したものである。【図3】 PSP1単離体PSP1-1、PSP1-  
2、PSP1-3およびPSP1-4の複数のcDNA  
配列を整列化したものである。50 【図4】 PSP1単離体PSP1-1、PSP1-  
2、PSP1-3およびPSP1-4の複数のcDNA  
配列を整列化したものである。

【図5】 PSP1単離体PSP1-1、PSP1-2、PSP1-3およびPSP1-4の複数のcDNA配列を整列化したものである。

【図6】 PSP1単離体PSP1-1、PSP1-2、PSP1-3およびPSP1-4の複数のcDNA配列を整列化したものである。

【図7】 PSP1単離体PSP1-1、PSP1-2、PSP1-3およびPSP1-4の複数のcDNA配列を整列化したものである。

【図8】 PSP1単離体PSP1-1、PSP1-

2、PSP1-3およびPSP1-4の複数のcDNA配列を整列化したものである。

【図9】 PSP1単離体PSP1-1、PSP1-2、PSP1-3およびPSP1-4の複数のcDNA配列を整列化したものである。

【図10】 PSP1単離体PSP1-1、PSP1-2、PSP1-3およびPSP1-4の複数のcDNA配列を整列化したものである。

【図11】 推定ヒトセリンプロテアーゼでPSP1-1のアミノ酸配列を整列化したものである。

【図1】

```

41 SGTSDFRVRVYGTPLSWARLSVGVTEPRACLTSGTPGPRAQLTAVTPDT 90
   . | . : . | : : : . : . | | . . . | . . . | . . . | . . .
2 KKTTLALSRLALSGLALSPLSATAAETSSATTAQQMPSLAPMLEKVMPS 51

91 RTREASENSGTRSRAWLAVALGAGGAVLLLLWGGRGPPAVLAAPVSPPPP 140
   . . . | . | . | . : : : . : . . . : : : . . . . . | | .
52 VVSINVEGSTTVNTPRMPRNFOQ.....FFGDD....SPFCQEGSPFQ 90

141 ASPRSQYNFIADVVEKTAPAVVYIEILDRHPFLGREVPISNGSGFVVAAD 190
   . | | : | . : : : . : . : : : . . . : | | . : . | |
91 SSPFCQ...GGQGGNGGGQQQKFMA.....GSGVIIDAD 122

191 .GLIVTNAHVVADRRRVRVRLLSGDTYEAVVTAVDPAVADIATLRIQTKEP 239
   | . : | | | . | | | . : : | . | . : : | . : : | | . | | . : . | | . .
123 KGYVVTNNHVVDNATVIKQLSDGRKFDAKMVGKDPKSDIALIQIQNPKN 172

240 LPTLPLGRSADVRQGEFVVAMGSPFALQNTITSGIVSSAQRPAIDLGLPQ 289
   | . : : : . | . : | | : : | : : | . | | : | | | | . . | . . | | .
173 LTAIKMADSDALRVGDYTVGIGNPFGLGETVTSGIVSALGRS....GLNA 218

290 TNVE.YIQTDAAIDFGNSGGPLVNLGDEVIGVNTMKVTA.....GISFAI 333
   . | | : | | | | | : | | | | : | | | : | | : | | . : . | | : | |
219 ENYENFIQTDAAINRGNSGGALVNLGELIGINTAILAPDGGNIGIGFAI 268

334 PSDRLREF.....LHRGE..... 346
   | | : : : : . : . | |
269 PSNMVKNLTSQMVEYGVKRGELGIMGTELNSELAKAMKVDAQRGAFFVSQ 318

347 .KKNSSSGISG.....SQRRIYGV.....MLTL... 369
   . | | | . : . | . : | . | . : | |
319 VLPNSSAAKAGIKAGDVITSLNGKPISSFAALRAQVGTMPVGSKLTLGLL 348

370 .....SPSILAEQLREPSPFDVQHGVLHVKVIL 398
   | . | | : : : . | | . : : | | : : . |
369 RDGKQVNVNLELQQSSQNQVDSSSIFNGIEGAEMSNKKGKQGVVNVNVT 418

399 GSPAHRAGLRPGDVILAIGEQMVQNAEDVYEAVRTQ.SQLAVQIRRGRET 447
   | . | | . | | : . | | | : . : | | . | . : : . . | | : . | |
419 GTPAAQIGLKKGDVIIGANQAVKNIAELRKVLDSKPSVLALNIQRGDRH 468

448 LTLYVTPEVTE 458
   | . : . . : . .
469 LPVNAVISLNP 479

```



## 【図2】

	1		50
PSP1-2	CGTGGATCCC	GAGAAAGAGG	CGCAGGACGA GGAGGCAGAA CCCGACTGGC
PSP1-1	CGTGGATCCC	GAGAAAGAGG	CGCAGGACGA GGAGGCAGAA CCCGACTGGC
PSP1-3	CGTGGATCCC	GAGAAAGAGG	CGCAGGACGA GGAGGCAGAA CCCGACTGGC
PSP1-4	CGTGGATCCC	GAGAAAGAGG	CGCAGGACGA GGAGGCAGAA CCCGACTGGC
	51		100
PSP1-2	GCGTAGAGCA	GCAGCACGAG	CAGTAGGAAG CAGTCACCCG GAAGCCTGGG
PSP1-1	GCGTAGAGCA	GCAGCACGAG	CAGTAGGAAG CAGTCACCCG GAAGCCTGGG
PSP1-3	GCGTAGAGCA	GCAGCACGAG	CAGTAGGAAG CAGTCACCCG GAAGCCTGGG
PSP1-4	GCGTAGAGCA	GCAGCACGAG	CAGTAGGAAG CAGTCACCCG GAAGCCTGGG
	101		150
PSP1-2	GGCGAGAGGC	GAAGTGGTCA	GGCGCCGAAG GCCGAGAGCA CGCGGGGATC
PSP1-1	GGCGAGAGGC	GAAGTGGTCA	GGCGCCGAAG GCCGAGAGCA CGCGGGGATC
PSP1-3	GGCGAGAGGC	GAAGTGGTCA	GGCGCCGAAG GCCGAGAGCA CGCGGGGATC
PSP1-4	GGCGAGAGGC	GAAGTGGTCA	GGCGCCGAAG GCCGAGAGCA CGCGGGGATC
	151		200
PSP1-2	GGTCTCTTCC	CGCCGGGTCT	CTTACCGGTG CGAGTCAAAG AGCCGCTCCG
PSP1-1	GGTCTCTTCC	CGCCGGGTCT	CTTACCGGTG CGAGTCAAAG AGCCGCTCCG
PSP1-3	GGTCTCTTCC	CGCCGGGTCT	CTTACCGGTG CGAGTCAAAG AGCCGCTCCG
PSP1-4	GGTCTCTTCC	CGCCGGGTCT	CTTACCGGTG CGAGTCAAAG AGCCGCTCCG
	201		250
PSP1-2	GCCCCGGCCC	TGAGGGAAGC	TCCATAACTG CTGCTTCAGG AGCGCCCGGC
PSP1-1	GCCCCGGCCC	TGAGGGAAGC	TCCATAACTG CTGCTTCAGG AGCGCCCGGC
PSP1-3	GCCCCGGCCC	TGAGGGAAGC	TCCATAACTG CTGCTTCAGG AGCGCCCGGC
PSP1-4	GCCCCGGCCC	TGAGGGAAGC	TCCATAACTG CTGCTTCAGG AGCGCCCGGC
	251		300
PSP1-2	CGTCGCCGCC	GCCGCCATTT	TCGCGCCCGG CCGCAGGGGC TCTTGGAAG
PSP1-1	CGTCGCCGCC	GCCGCCATTT	TCGCGCCCGG CCGCAGGGGC TCTTGGAAG
PSP1-3	CGTCGCCGCC	GCCGCCATTT	TCGCGCCCGG CCGCAGGGGC TCTTGGAAG
PSP1-4	CGTCGCCGCC	GCCGCCATTT	TCGCGCCCGG CCGCAGGGGC TCTTGGAAG

## 【図3】

## 図2の続き

	301		350
PSP1-2	GCGGAGTCTT	TGGGCATCCG	CCCGGGGTGA GGGGACCCGA AGTCCTGAGG
PSP1-1	GCGGAGTCTT	TGGGCATCCG	CCCGGGGTGA GGGGACCCGA AGTCCTGAGG
PSP1-3	GCGGAGTCTT	TGGGCATCCG	CCCGGGGTGA GGGGACCCGA AGTCCTGAGG
PSP1-4	GCGGAGTCTT	TGGGCATCCG	CCCGGGGTGA GGGGACCCGA AGTCCTGAGG
	351		400
PSP1-2	CGCGCCGGAA	GGGCTAGCGG	TCCCAGCATA CCCCGCGGCC CCTTGGGCCG
PSP1-1	CGCGCCGGAA	GGGCTAGCGG	TCCCAGCATA CCCCGCGGCC CCTTGGGCCG
PSP1-3	CGCGCCGGAA	GGGCTAGCGG	TCCCAGCATA CCCCGCGGCC CCTTGGGCCG
PSP1-4	CGCGCCGGAA	GGGCTAGCGG	TCCCAGCATA CCCCGCGGCC CCTTGGGCCG
	401		450
PSP1-2	TCTCACAAC	TCTCACAAC	CGCGTCCGGC GGAGACCACA ATTCCCGGCA TTCGTGGGGC
PSP1-1	TCTCACAAC	TCTCACAAC	CGCGTCCGGC GGAGACCACA ATTCCCGGCA TTCGTGGGGC
PSP1-3	TCTCACAAC	TCTCACAAC	CGCGTCCGGC GGAGACCACA ATTCCCGGCA TTCGTGGGGC
PSP1-4	TCTCACAAC	TCTCACAAC	CGCGTCCGGC GGAGACCACA ATTCCCGGCA TTCGTGGGGC
	451		500
PSP1-2	AGGGAGGAGT	CGGCCTCCCG	GAATCCTGGT CCCGGCGTGC ACTTCTGAAG
PSP1-1	AGGGAGGAGT	CGGCCTCCCG	GAATCCTGGT CCCGGCGTGC ACTTCTGAAG
PSP1-3	AGGGAGGAGT	CGGCCTCCCG	GAATCCTGGT CCCGGCGTGC ACTTCTGAAG
PSP1-4	AGGGAGGAGT	CGGCCTCCCG	GAATCCTGGT CCCGGCGTGC ACTTCTGAAG
	501		550
PSP1-2	GACTTCAGGT	ACCGGCGTGC	CCCGCGTCCT ACTGTCCGCC TGCTCGCGTC
PSP1-1	GACTTCAGGT	ACCGGCGTGC	CCCGCGTCCT ACTGTCCGCC TGCTCGCGTC
PSP1-3	GACTTCAGGT	ACCGGCGTGC	CCCGCGTCCT ACTGTCCGCC TGCTCGCGTC
PSP1-4	GACTTCAGGT	ACCGGCGTGC	CCCGCGTCCT ACTGTCCGCC TGCTCGCGTC
	551		600
PSP1-2	CTGGGTGCCG	CCTCTGAGTA	GGGCGGGCGA GGAGGCAGCC AAGGCGGAGC
PSP1-1	CTGGGTGCCG	CCTCTGAGTA	GGGCGGGCGA GGAGGCAGCC AAGGCGGAGC
PSP1-3	CTGGGTGCCG	CCTCTGAGTA	GGGCGGGCGA GGAGGCAGCC AAGGCGGAGC
PSP1-4	CTGGGTGCCG	CCTCTGAGTA	GGGCGGGCGA GGAGGCAGCC AAGGCGGAGC

【図4】

## 図3の続き

	601		650
PSP1-2	TGATGGCTGC	GCCGAGGGCG	GGGCGGGGTG CAGGCTGGAG CCTTCGGGCA
PSP1-1	TGATGGCTGC	GCCGAGGGCG	GGGCGGGGTG CAGGCTGGAG CCTTCGGGCA
PSP1-3	TGATGGCTGC	GCCGAGGGCG	GGGCGGGGTG CAGGCTGGAG CCTTCGGGCA
PSP1-4	TGATGGCTGC	GCCGAGGGCG	GGGCGGGGTG CAGGCTGGAG CCTTCGGGCA
	651		700
PSP1-2	TGGCGGGCTT	TGGGGGGCAT	TTGCTGGGGG AGGAGACCCC GTTTGACCCC
PSP1-1	TGGCGGGCTT	TGGGGGGCAT	TCGCTGGGGG AGGAGACCCC GTTTGACCCC
PSP1-3	TGGCGGGCTT	TGGGGGGCAT	TCGCTGGGGG AGGAGACCCC GTTTGACCCC
PSP1-4	TGGCGGGCTT	TGGGGGGCAT	TCGCTGGGGG AGGAGACCCC GTTTGACCCC
	701		750
PSP1-2	TGACCTCCGG	GCCCTGCTGA	CGTCAGGAAC TTCTGACCCC CGGGCCCGAG
PSP1-1	TGACCTCCGG	GCCCTGCTGA	CGTCAGGAAC TTCTGACCCC CGGGCCCGAG
PSP1-3	TGACCTCCGG	GCCCTGCTGA	CGTCAGGAAC TTCTGACCCC CGGGCCCGAG
PSP1-4	TGACCTCCGG	GCCCTGCTGA	CGTCAGGAAC TTCTGACCCC CGGGCCCGAG
	751		800
PSP1-2	TGACTTATGG	GACCCCCAGT	CTCTGGGCCC GGTGTCTGT TGGGGTCACT
PSP1-1	TGACTTATGG	GACCCCCAGT	CTCTGGGCCC GGTGTCTGT TGGGGTCACT
PSP1-3	TGACTTATGG	GACCCCCAGT	CTCTGGGCCC GGTGTCTGT TGGGGTCACT
PSP1-4	TGACTTATGG	GACCCCCAGT	CTCTGGGCCC GGTGTCTGT TGGGGTCACT
	801		850
PSP1-2	GAACCCCGAG	CATGCCTGAC	GTCTGGGACC CCGGGTCCCC GGGCACAAC
PSP1-1	GAACCCCGAG	CATGCCTGAC	GTCTGGGACC CCGGGTCCCC GGGCACAAC
PSP1-3	GAACCCCGAG	CATGCCTGAC	GTCTGGGACC CCGGGTCCCC GGGCACAAC
PSP1-4	GAACCCCGAG	CATGCCTGAC	GTCTGGGACC CCGGGTCCCC GGGCACAAC
	851		900
PSP1-2	GACTGCGGTG	ACCCAGATA	CCAGGACCCG GGAGGCCTCA GAGAACTCTG
PSP1-1	GACTGCGGTG	ACCCAGATA	CCAGGACCCG GGAGGCCTCA GAGAACTCTG
PSP1-3	GACTGCGGTG	ACCCAGATA	CCAGGACCCG GGAGGCCTCA GAGAACTCTG
PSP1-4	GACTGCGGTG	ACCCAGATA	CCAGGACCCG GGAGGCCTCA GAGAACTCTG

## 【図5】

## 図4の続き

	901		950
PSP1-2	GAACCCGTTT	GCGCGCGTGG	CTGGCGGTGG CGCTGGGCGC TGGGGGGGCA
PSP1-1	GAACCCGTTT	GCGCGCGTGG	CTGGCGGTGG CGCTGGGCGC TGGGGGGGCA
PSP1-3	GAACCCGTTT	GCGCGCGTGG	CTGGCGGTGG CGCTGGGCGC TGGGGGGGCA
PSP1-4	GAACCCGTTT	GCGCGCGTGG	CTGGCGGTGG CGCTGGGCGC TGGGGGGGCA
	951		1000
PSP1-2	GTGCTGTTGT	TGTTGTGGGG	CGGGGGTCGG GGTCTCTCCG CCGTCTCTCG
PSP1-1	GTGCTGTTGT	TGTTGTGGGG	CGGGGGTCGG GGTCTCTCCG CCGTCTCTCG
PSP1-3	GTGCTGTTGT	TGTTGTGGGG	CGGGGGTCGG GGTCTCTCCG CCGTCTCTCG
PSP1-4	GTGCTGTTGT	TGTTGTGGGG	CGGGGGTCGG GGTCTCTCCG CCGTCTCTCG
	1001		1050
PSP1-2	CGCCGTCCCT	AGCCCGCCGC	CCGCTTCTCC CCGGAGTCAG TACAATTCA
PSP1-1	CGCCGTCCCT	AGCCCGCCGC	CCGCTTCTCC CCGGAGTCAG TACAATTCA
PSP1-3	CGCCGTCCCT	AGCCCGCCGC	CCGCTTCTCC CCGGAGTCAG TACAATTCA
PSP1-4	CGCCGTCCCT	AGCCCGCCGC	CCGCTTCTCC CCGGAGTCAG TACAATTCA
	1051		1100
PSP1-2	TCGCAGATGT	GGTGGAGAAG	ACAGCACCTG CCGTGGTCTA TATCGAGATC
PSP1-1	TCGCAGATGT	GGTGGAGAAG	ACAGCACCTG CCGTGGTCTA TATCGAGATC
PSP1-3	TCGCAGATGT	GGTGGAGAAG	ACAGCACCTG CCGTGGTCTA TATCGAGATC
PSP1-4	TCGCAGATGT	GGTGGAGAAG	ACAGCACCTG CCGTGGTCTA TATCGAGATC
	1101		1150
PSP1-2	CTGGACCGGC	ACCCTTTCTT	GGGCCGCGAG GTCCCTATCT CGAACGGGCTC
PSP1-1	CTGGACCGGC	ACCCTTTCTT	GGGCCGCGAG GTCCCTATCT CGAACGGGCTC
PSP1-3	CTGGACCGGC	ACCCTTTCTT	GGGCCGCGAG GTCCCTATCT CGAACGGGCTC
PSP1-4	CTGGACCGGC	ACCCTTTCTT	GGGCCGCGAG GTCCCTATCT CGAACGGGCTC
	1151		1200
PSP1-2	AGGATTCGTG	GTGGCTGCCG	ATGGGCTCAT TGTACCAAC GCCCATGTGG
PSP1-1	AGGATTCGTG	GTGGCTGCCG	ATGGGCTCAT TGTACCAAC GCCCATGTGG
PSP1-3	AGGATTCGTG	GTGGCTGCCG	ATGGGCTCAT TGTACCAAC GCCCATGTGG
PSP1-4	AGGATTCGTG	GTGGCTGCCG	ATGGGCTCAT TGTACCAAC GCCCATGTGG

## 【図6】

## 図5の続き

	1201		1250
PSP1-2	TGGCTGATCG GCGCAGAGTC CGTGTGAGAC TGCTAAGCGG CGACACGTAT		
PSP1-1	TGGCTGATCG GCGCAGAGTC CGTGTGAGAC TGCTAAGCGG CGACACGTAT		
PSP1-3	TGGCTGATCG GCGCAGAGTC CGTGTGAGAC TGCTAAGCGG CGACACGTAT		
PSP1-4	TGGCTGATCG GCGCAGAGTC CGTGTGAGAC TGCTAAGCGG CGACACGTAT		
	1251		1300
PSP1-2	GAGGCCGTGG TCACAGCTGT GGATCCCGTG GCAGACATCG CAACGCTGAG		
PSP1-1	GAGGCCGTGG TCACAGCTGT GGATCCCGTG GCAGACATCG CAACGCTGAG		
PSP1-3	GAGGCCGTGG TCACAGCTGT GGATCCCGTG GCAGACATCG CAACGCTGAG		
PSP1-4	GAGGCCGTGG TCACAGCTGT GGATCCCGTG GCAGACATCG CAACGCTGAG		
	1301		1350
PSP1-2	GATTCAGACT AAGGAGCCTC TCCCCACGCT GCCTCTGGGA CGCTCAGCTG		
PSP1-1	GATTCAGACT AAGGAGCCTC TCCCCACGCT GCCTCTGGGA CGCTCAGCTG		
PSP1-3	GATTCAGACT AAGGAGCCTC TCCCCACGCT GCCTCTGGGA CGCTCAGCTG		
PSP1-4	GATTCAGACT AAGGAGCCTC TCCCCACGCT GCCTCTGGGA CGCTCAGCTG		
	1351		1400
PSP1-2	ATGTCCGGCA AGGGGAGTTT GTTGTGCCA TGGGAAGTCC CTTTGCACTG		
PSP1-1	ATGTCCGGCA AGGGGAGTTT GTTGTGCCA TGGGAAGTCC CTTTGCACTG		
PSP1-3	ATGTCCGGCA AGGGGAGTTT GTTGTGCCA TGGGAAGTCC CTTTGCACTG		
PSP1-4	ATGTCCGGCA AGGGGAGTTT GTTGTGCCA TGGGAAGTCC CTTTGCACTG		
	1401		1450
PSP1-2	CAGAACACGA TCACATCCGG CATTGTTAGC TCTGCTCAGC GTCCAGCCAG		
PSP1-1	CAGAACACGA TCACATCCGG CATTGTTAGC TCTGCTCAGC GTCCAGCCAG		
PSP1-3	CAGAACACGA TCACATCCGG CATTGTTAGC TCTGCTCAGC GTCCAGCCAG		
PSP1-4	CAGAACACGA TCACATCCGG CATTGTTAGC TCTGCTCAGC GTCCAGCCAG		
	1451		1500
PSP1-2	AGACCTGGGA CTCCCCAAA CCAATGTGGA ATACATTCAA ACTGATGCAG		
PSP1-1	AGACCTGGGA CTCCCCAAA CCAATGTGGA ATACATTCAA ACTGATGCAG		
PSP1-3	AGACCTGGGA CTCCCCAAA CCAATGTGGA ATACATTCAA ACTGATGCAG		
PSP1-4	AGACCTGGGA CTCCCCAAA CCAATGTGGA ATACATTCAA ACTGATGCAG		

## 【図7】

## 図6の続き

	1501		1550
PSP1-2	CTATTGATTT	TGGAACTCT	GGAGGTCCCC TGGTTAACCT .....
PSP1-1	CTATTGATTT	TGGAACTCT	GGAGGTCCCC TGGTTAACCT .....
PSP1-3	CTATTGATTT	TGGAACTCT	GGAGGTCCCC TGGTTAACCT GGTGAGTGAG
PSP1-4	CTATTGATTT	TGGAACTCT	GGAGGTCCCC TGGTTAACCT .....
	1551		1600
PSP1-2	.....	.....	.....
PSP1-1	.....	.....	.....
PSP1-3	ACATCCTTCC	TTCCAAGAAT	CCCTGCCCCA GGTCAGTGTG GGAAGGGTAG
PSP1-4	.....	.....	.....
	1601		1650
PSP1-2	.....	.....	.....
PSP1-1	.....	.....	.....
PSP1-3	GTTTCCCCTA	ATTCAAGGAT	GTTTGGTCAA GTTTCTGAGC AGTTCTTTGT
PSP1-4	.....	.....	.....
	1651		1700
PSP1-2	.....	.....	.....
PSP1-1	.....	.....	.....
PSP1-3	TGGCTATCTC	TCAATATCCA	ACCAGATCTC CCCAACACTT GCTGGTACTT
PSP1-4	.....	.....	.....
	1701		1750
PSP1-2	.....	.....	.....
PSP1-1	.....	.....	.....
PSP1-3	TTGTTCCGGT	GCCCCCATCC	CCTACTATTT GTTTAGGCTA GGGAACTGGG
PSP1-4	.....	.....	.....GGCTA GGGAACTGGG
	1751		1800
PSP1-2	.....	.....GGATG	GGGAGGTGAT TGGAGTGAAC ACCATGAAGG
PSP1-1	.....	.....GGATG	GGGAGGTGAT TGGAGTGAAC ACCATGAAGG
PSP1-3	GGCTGTATCC	CTGCAGGATG	GGGAGGTGAT TGGAGTGAAC ACCATGAAGG
PSP1-4	GGCTGTATCC	CTGCAGGATG	GGGAGGTGAT TGGAGTGAAC ACCATGAAGG

【図8】

## 図7の続き

	1801		1850
PSP1-2	TCACAGCTGG AATCTCCTTT GCCATCCCTT CTGATCGTCT TCGAGAGTTT		
PSP1-1	TCACAGCTGG AATCTCCTTT GCCATCCCTT CTGATCGTCT TCGAGAGTTT		
PSP1-3	TCACAGCTGG AATCTCCTTT GCCATCCCTT CTGATCGTCT TCGAGAGTTT		
PSP1-4	TCACAGCTGG AATCTCCTTT GCCATCCCTT CTGATCGTCT TCGAGAGTTT		
	1851		1900
PSP1-2	CTGCATCGTG GGGAAAAGAA GAATTCCTCC TCCGGAATCA GTGGGTCCCA		
PSP1-1	CTGCATCGTG GGGAAAAGAA GAATTCCTCC TCCGGAATCA GTGGGTCCCA		
PSP1-3	CTGCATCGTG GGGAAAAGAA GAATTCCTCC TCCGGAATCA GTGGGTCCCA		
PSP1-4	CTGCATCGTG GGGAAAAGAA GAATTCCTCC TCCGGAATCA GTGGGTCCCA		
	1901		1950
PSP1-2	GCGGCGCTAC ATTGGGGTGA TGATGCTGAC CCTGAGTCCC AGCATCCTTG		
PSP1-1	GCGGCGCTAC ATTGGGGTGA TGATGCTGAC CCTGAGTCCC AGCATCCTTG		
PSP1-3	GCGGCGCTAC ATTGGGGTGA TGATGCTGAC CCTGAGTCCC AGCATCCTTG		
PSP1-4	GCGGCGCTAC ATTGGGGTGA TGATGCTGAC CCTGAGTCCC A.....		
	1951		2000
PSP1-2	CTGAACTACA GCTTCGAGAA CCAAGCTTTC CCGATGTTCA GCATGGTGTA		
PSP1-1	CTGAACTACA GCTTCGAGAA CCAAGCTTTC CCGATGTTCA GCATGGTGTA		
PSP1-3	CTGAACTACA GCTTCGAGAA CCAAGCTTTC CCGATGTTCA GCATGGTGTA		
PSP1-4	.....		
	2001		2050
PSP1-2	CTCATCCATA AAGTCATCCT GGGCTCCCCT GCACACCGGG CTGGTCTGCG		
PSP1-1	CTCATCCATA AAGTCATCCT GGGCTCCCCT GCACACCGGG CTGGTCTGCG		
PSP1-3	CTCATCCATA AAGTCATCCT GGGCTCCCCT GCACACCGGG CTGGTCTGCG		
PSP1-4	.....GGG CTGGTCTGCG		
	2051		2100
PSP1-2	GCCTGGTGAT GTGATTTTGG CCATTGGGGA GCAGATGGTA CAAAATGCTG		
PSP1-1	GCCTGGTGAT GTGATTTTGG CCATTGGGGA GCAGATGGTA CAAAATGCTG		
PSP1-3	GCCTGGTGAT GTGATTTTGG CCATTGGGGA GCAGATGGTA CAAAATGCTG		
PSP1-4	GCCTGGTGAT GTGATTTTGG CCATTGGGGA GCAGATGGTA CAAAATGCTG		

## 【図9】

## 図8の続き

	2101		2150
PSP1-2	AAGATGTTTA TGAAGCTGTT CGAACCCAAT CCCAGTTGGC AGTGCAGATC		
PSP1-1	AAGATGTTTA TGAAGCTGTT CGAACCCAAT CCCAGTTGGC AGTGCAGATC		
PSP1-3	AAGATGTTTA TGAAGCTGTT CGAACCCAAT CCCAGTTGGC AGTGCAGATC		
PSP1-4	AAGATGTTTA TGAAGCTGTT CGAACCCAAT CCCAGTTGGC AGTGCAGATC		
	2151		2200
PSP1-2	CGGCGGGGAC GAGAAACACT GACCTTATAT GTGACCCCTG AGGTCACAGA		
PSP1-1	CGGCGGGGAC GAGAAACACT GACCTTATAT GTGACCCCTG AGGTCACAGA		
PSP1-3	CGGCGGGGAC GAGAAACACT GACCTTATAT GTGACCCCTG AGGTCACAGA		
PSP1-4	CGGCGGGGAC GAGAAACACT GACCTTATAT GTGACCCCTG AGGTCACAGA		
	2201		2250
PSP1-2	ATGAATAGAT CACCAAGAGT ATGAGGCTCC TGCTCTGATT TCCTCCTTGC		
PSP1-1	ATGAATAGAT CACCAAGAGT ATGAGGCTCC TGCTCTGATT TCCTCCTTGC		
PSP1-3	ATGAATAGAT CACCAAGAGT ATGAGGCTCC TGCTCTGATT TCCTCCTTGC		
PSP1-4	ATGAATAGAT CACCAAGAGT ATGAGGCTCC TGCTCTGATT TCCTCCTTGC		
	2251		2300
PSP1-2	CTTTCTGGCT GAGGTTCTGA GGGCACCGAG ACAGAGGGTT AAATGAACCA		
PSP1-1	CTTTCTGGCT GAGGTTCTGA GGGCACCGAG ACAGAGGGTT AAATGAACCA		
PSP1-3	CTTTCTGGCT GAGGTTCTGA GGGCACCGAG ACAGAGGGTT AAATGAACCA		
PSP1-4	CTTTCTGGCT GAGGTTCTGA GGGCACCGAG ACAGAGGGTT AAATGAACCA		
	2301		2350
PSP1-2	GTGGGGGCAG GTCCCTCCAA CCACCAGCAC TGACTCCTGG GCTCTGAAGA		
PSP1-1	GTGGGGGCAG GTCCCTCCAA CCACCAGCAC TGACTCCTGG GCTCTGAAGA		
PSP1-3	GTGGGGGCAG GTCCCTCCAA CCACCAGCAC TGACTCCTGG GCTCTGAAGA		
PSP1-4	GTGGGGGCAG GTCCCTCCAA CCACCAGCAC TGACTCCTGG GCTCTGAAGA		
	2351		2400
PSP1-2	ATCACAGAAA CACTTTTAT ATAAAATAAA ATTATACCTA GCAACATAAA		
PSP1-1	ATCACAGAAA CACTTTTAT ATAAAATAAA ATTATACCTA GCAACATAAA		
PSP1-3	ATCACAGAAA CACTTTTAT ATAAAATAAA ATTATACCTA GCAACATAAA		
PSP1-4	ATCACAGAAA CACTTTTAT ATAAAATAAA ATTATACCTA GCAACATAAA		



## 【図10】

## 図9の続き

	2401		2450
PSP1-2	AAAAAAAAAA AA.....		
PSP1-1	AAAAAAAAAA AA.....		
PSP1-3	ATAGTAAAAA ATGAGGTGGG AGGGCTGGAT CTTTCCCCC ACCAAAAGGC		
PSP1-4	AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAA.....		
	2451		2500
PSP1-2	.....		
PSP1-1	.....		
PSP1-3	TAGAGGTAAA GCTGTATCCC CCTAAACTTA GGGGAGATAC TGGAGCTGAC		
PSP1-4	.....		
	2501		2550
PSP1-2	.....		
PSP1-1	.....		
PSP1-3	CATCCTGACC TCCTATTAAA GAAAATGAGC TGCTGAAAAA AAAAAAAAAA		
PSP1-4	.....		
	2551		
PSP1-2	.		
PSP1-1	.		
PSP1-3	A		
PSP1-4	.		

```

1 MAAP.....RAGRGAGWSLRAWRALGGIRWGRRLPTDLRALLTSSTGTS 44
   :||| |||:|. | . :. |:|. .. 
16 LAAPASQSLSRAGRSAPL.....AAGCPDRCEPARCPPQ.....PEHCE 54
   .....
45 DPRARVTYGTPLSWARLSVGVTETPRACLTSGTPGPRAQLTAV.....TP 88
   :||| |:|... :|.|. || :|. : .| :. 
55 GGRARDACGCCCEV.....CGAPEGAACGLQEGPCGEGLQCVPFPGVPASA 99
   .....
89 DTRTREASENSGTRSRAWLAVALCAGGAVLLLLWGWW.....RGPPAV 131
   .| |... :. |. .... :. | | |::: |.| |
100 TVRRRAQAGLCVCASSEPVCGSDANTYANLCQLRAASRRSERLHRPPVIV 149
   .....
132 LAAVPSPP....PASPRSQYNFIADVVEKTAPAVVYIEILDRHPFLGREV 177
   |. .... |. | .|||||||.|||||.||:: : || |||
150 LQRGACGGQGEDPNSLRHKYNFIAVDVEKIAPAVVHIELFRKLFPFSKREV 199
   .....
178 PISNGSGFVVAADGLIVTNAHVVAODRRVRVRLLSGOTYEAVVTAVDPVA 227
   |:..||||:|. .|||||||.||::|:| | .|.|||| :..|. |
200 PVASGSGFIVSEDGLIVTNAHVVTNKHRVKVELKNGATYEAKIKDVDEKA 249
   .....
228 DIATLRIQTKEPLTPLGRSADVRQGEFVVMAMSPFALQNTITSGIVSS 277
   ||| ::|: .:.|.| ||||::|.|||||:|:-|||:|.||||.
250 DIALIKIDHQGLKPVLILLGRSSELRPGFEFVVAIGSPFSLQNTVTTTGIVST 299
   .....
278 AQRPARDLGLPQTNVEYIQTDAADIFGNSGGPLVNLDGEVIGVNTMKVTA 327
   .||:::|||...:::||||| |::|:|||||||:|:|:|:|:|
300 TQRGGKELGLRNSOMDIYIQTDAIINYGNSGGPLVNLDGEVIGINTLKVTA 349
   .....
328 GISFAIPSDRLREFLHRGEKKNSSSGISGSQRRYIGVMMTLSPSILAEL 377
   |||||:::|.| :... ..| . .::|:|:|.|.|. | |
350 GISFAIPSDKIKKFLTESHDRQ.AKGKAITKKKYIGIRMMSLTSSKAKEL 398
   .....
378 QLREPSFPDVBQHGVLHKVILGPSAHRAGLRPGDVILAIGEQMVSQNAEDV 427
   . |...||| |. | .|| :|. :|:..|:|:|.|. | | .|:|
399 KDRHRDFPDVISGAYIIEVIPDTPAEAGGLKENDVIISINGQS SVVSANDV 448
   .....
428 YEAVRTQS QLA VQIRRGRETTLTVTPEVTE 458
   ::::| |.: :|||.|. :|. || :.
449 SDVIKRESTLN MVVRGNEDIMITVIPEEID 479

```

(51) Int. Cl. <sup>a</sup>

FI

C 1 2 P 21/02

C 1 2 Q 1/37

1/68

A 6 1 K 37/54

$$z$$

ADS

(71)出願人 591002957  
スミスクライン・ビーチャム・コーポレイ  
ション  
SMITHKLINE BEECHAM  
CORPORATION  
アメリカ合衆国ペンシルベニア州19406-  
0939、キング・オブ・プルシア、スウェー  
ドランド・ロード709番

(72)発明者 エリック・ハワード・カラン  
イギリス、シーエム19・5エイダブリュ  
ー、エセックス、ハーロウ、ニュー・フロ  
ンティアーズ・サイエンス・パーク・サウ  
ス、スミスクライン・ビーチャム・ファーマ  
シューティカルズ

(72)発明者 ヘレン・エリザベス・クリンケンベアード  
イギリス、シーエム19・5エイダブリュ  
ー、エセックス、ハーロウ、ニュー・フロ  
ンティアーズ・サイエンス・パーク・サウ  
ス、スミスクライン・ビーチャム・ファーマ  
シューティカルズ

(72)発明者 マイケル・ジョゼフ・ブラウン  
イギリス、シーエム19・5エイダブリュ  
ー、エセックス、ハーロウ、ニュー・フロ  
ンティアーズ・サイエンス・パーク・サウ  
ス、スミスクライン・ビーチャム・ファーマ  
シューティカルズ

(72)発明者 クリストファー・デイビッド・サウザン  
イギリス、シーエム19・5エイダブリュ  
ー、エセックス、ハーロウ、ニュー・フロ  
ンティアーズ・サイエンス・パーク・サウ  
ス、スミスクライン・ビーチャム・ファーマ  
シューティカルズ

(72)発明者 カレサ・エル・クリーシー  
アメリカ合衆国19406ペンシルベニア州キ  
ング・オブ・プルシア、スウェードラン  
ド・ロード709番 スミスクライン・ビー  
チャム・ファーマシューティカルズ

(72)発明者 ジョージ・ビー・リビ  
アメリカ合衆国19406ペンシルベニア州キ  
ング・オブ・プルシア、スウェードラン  
ド・ロード709番 スミスクライン・ビー  
チャム・ファーマシューティカルズ

【外国語明細書】

## HUMAN SERINE PROTEASE

The present invention relates to isolated human serine protease (PSP1) polynucleotides, their homologs and isoforms and polymorphic variants and their detection;  
5 to essentially pure PSP1 proteins; and to compositions and methods of producing and using PSP1 polynucleotides and proteins.

Mutations in the presenilins (PS-1 and PS-2) account for ~95% (75% and 20%, respectively) of all cases of early onset familial Alzheimer's disease (FAD). See R. Sherrington *et al.*, *Nature* 375, 754-760 (1995); E.I. Rogaev *et al.*, *Nature* 376, 775-778  
10 (1995); and E. Levy-Lahad *et al.*, *Science* 269, 973-977 (1995). The presenilins are highly homologous (67% identical), multi-membrane spanning proteins whose function is unknown.

It has been demonstrated that the 46 kDa full-length PS-1 protein is normally processed to 28 kDa and 18 kDa fragments; PS-2 has been reported to be similarly cleaved.  
15 See M. Mercken *et al.*, *FEBS Letters* 389, 297-303 (1996). The predicted cleavage site(s) to account for fragments of this size would be in a region of the protein coded for by exon 8 and exon 9. Exon 8 is a hot spot for mutations leading to FAD. Thus, this region of PS-1, and potentially the cleavage of PS-1 in this region by a presenilinase protease, are important events in the functionality of the protein. A region of PS-1 spanning exons 8-11.  
20 has been demonstrated in the present invention to specifically bind a protease, PSP1, whose activity against its endogenous substrates and/or ability to bind to PS-1 are important in the pathology of neurodegeneration associated with AD, frontal lobe dementia, cortical lewy body disease, dementia of parkinson's disease, acute and chronic phases of degeneration following stroke or head injury, neuronal degeneration found in motor neurone disease,  
25 AIDS dementia and chronic epileps. Thus, a need exists for provision of the nucleotide and amino acid sequences corresponding to PSP1, for modulators of PSP1 binding to PS-1, and/or modulators of PSP1's proteolytic activity, for methods to identify such modulators and for reagents useful in such methods.

Accordingly, one aspect of the present invention is an isolated polynucleotide  
30 encoding a biologically active PSP1 polypeptide.

Another aspect of the invention is an isolated polynucleotide selected from the group consisting of:

整理番号 156220

(a) a polynucleotide encoding PSP1-1 having the nucleotide sequence as set forth in SEQ ID NO: 24 from nucleotide 603 to 1979; and

(b) a polynucleotide substantially similar to SEQ ID NO: 24.

Another aspect of the invention is an isolated polynucleotide selected from the group consisting of:

5 (a) a polynucleotide encoding PSP1-2 having the nucleotide sequence as set forth in SEQ ID NO: 23 from nucleotide 603 to 1979; and

(b) a polynucleotide substantially similar to SEQ ID NO: 23.

Another aspect of the invention is an isolated polynucleotide selected from the group consisting of:

10 (a) a polynucleotide encoding PSP1-3 having the nucleotide sequence as set forth in SEQ ID NO: 26 from nucleotide 603 to 1736; and

(b) a polynucleotide substantially similar to SEQ ID NO: 26.

Another aspect of the invention is an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide encoding PSP1-4 having the nucleotide sequence as set forth in SEQ ID NO: 28 from nucleotide 603 to 1913; and

(b) a polynucleotide substantially similar to SEQ ID NO: 28.

15 In a further aspect the invention provides any isolated polynucleotide as above defined wherein nucleotides 672 and 1435 are independently selected from C and T, hereinafter referred to as 'polymorphic variants'.

Another aspect of the invention is the functional polypeptides encoded by the polynucleotides of the invention.

20 Another aspect of the invention is an antisense oligonucleotide comprising a sequence which is capable of binding to the polynucleotides of the invention or D87258.

Another aspect of the invention is modulators of the polypeptides of the invention or of D87258.

25 Another aspect of the invention is a method for assaying a medium for the presence of a substance that modulates PSP1 or D87258 activity by affecting the binding of PSP1 or D87258 to cellular binding partners comprising the steps of:

(a) providing a PSP1 or D87258 protein having the amino acid sequence of PSP1-1, PSP1-2, PSP1-3 or PSP1-4 or D87258, or a functional derivative or polymorphic variant thereof and a cellular binding partner or synthetic analog thereof;

整理番号 156220

(b) incubating with a test substance which is suspected of modulating PSP1 or D87258 activity under conditions which permit the formation of a PSP1 or D87258 protein/cellular binding partner complex;

(c) assaying for the presence of the complex, free PSP1 or D87258 protein or free cellular binding partner; and

(d) comparing to a control to determine the effect of the substance.

Another aspect of the invention is a method for assaying a medium for the presence of a substance that modulates PSP1 or D87258 activity by inhibiting proteolytic activity on a cellular substrate comprising the steps of:

(a) providing a PSP1 or D87258 protein having the amino acid sequence of PSP1-1, PSP1-2, PSP1-3 or PSP1-4 or D87258, or a functional fragment or polymorphic variant thereof and a cellular substrate or synthetic analog thereof;

(b) incubating with a test substance which is suspected of inhibiting PSP1 or D87258 activity under conditions which permit the formation of a PSP1 enzyme/substrate complex and subsequent cleavage of the substrate;

(c) assaying for the presence of proteolytically cleaved substrate; and

(d) comparing to a control to determine the effect of the substance.

Another aspect of the invention is a method for assaying for the presence of a substance that modulates PSP1 or D87258 activity by direct binding to PSP1 or D87258 protein comprising the steps of:

(a) providing a labelled PSP1 or D87258 protein having the amino acid sequence of PSP1-1, PSP1-2, PSP1-3 or PSP1-4 or D87258 or a functional derivative or polymorphic variant thereof;

(b) providing solid support-associated modulator candidates;

(c) incubating a mixture of the labelled PSP1 or D87258 protein with the support-associated modulator candidates under conditions which can permit the formation of a PSP1 protein/modulator candidate complex;

(d) separating the solid support from free soluble labelled PSP1 or D87258 protein;

(e) assaying for the presence of solid support-associated labelled protein;

(f) isolating the solid support complexed with labelled PSP1 or D87258 protein; and

(g) identifying the modulator candidate.

整理番号 156220

Another aspect of the invention is PSP1 or D87258 protein modulating compounds identified by the methods of the invention.

Another aspect of the invention is a method for the treatment of a patient having need to modulate PSP1 or D87258 activity comprising administering to the patient a  
5 therapeutically effective amount of the modulating compounds of the invention.

Another aspect of the invention is a method of diagnosing conditions associated with PSP1 or D87258 protein deficiency which comprises:

- (a) isolating a polynucleotide sample from an individual;
- (b) assaying the polynucleotide sample and a polynucleotide of the invention  
10 encoding PSP1 or D87258; and
- (c) comparing differences between the polynucleotide sample and the PSP1 or D87258 polynucleotide, wherein any differences indicate mutations in the PSP1 or D87258 sequence.

Another aspect of the invention is a method of treating conditions which are related  
15 to insufficient PSP1 or D87258 protein function which comprises:

- (a) isolating cells from a patient deficient in PSP1 or D87258 protein function;
- (b) altering the cells by transfecting the polynucleotide of the invention or D87258 into the cells wherein a PSP1 or D87258 protein is expressed; and
- (c) introducing the cells back to the patient to alleviate the condition.

Another aspect of the invention is a method of treating conditions which are related  
20 to insufficient PSP1 or D87258 protein function which comprises administering the polynucleotide of the invention to a patient deficient in PSP1 protein function wherein a PSP1 or D87258 protein is expressed and alleviates the condition.

Another aspect of the invention is an antibody immunoreactive with PSP1 or  
25 D87258 or an immunogen thereof.

Another aspect of the invention is a transgenic non-human animal capable of expressing in any cell thereof the polynucleotide of the invention.

Another aspect of the invention is a method for determining the genetic predisposition to neurodegeneration in a patient comprising detecting PSP1 or D87258 polymorphisms in a sample from a patient. Yet another aspect of the invention is

isolated polynucleotide having the nucleotide sequence as set forth in SEQ ID NO: 32, 33,  
34, 35, 36, 37, 38, 39, or 40.

Figure 1 is an amino acid sequence alignment of PSP1-1 with *E. coli* htrA.

整理番号 156220

Figure 2 is a multiple cDNA sequence alignment of the PSP1 isolates *PSP1-1*, *PSP1-2*, *PSP1-3* and *PSP1-4*.

Figure 3 is an amino acid sequence alignment of PSP1-1 with a putative human serine protease.

5 As used herein, the term "PSP1 polynucleotide" or "*PSP1*" refers to DNA molecules comprising a nucleotide sequence that encodes PSP1 and alternative splice variants, i.e., homologs and isoforms, and polymorphic variants. PSP1 binds to a region encompassing amino acids 269-413 of the human PS-1 protein, contains a conserved serine protease motif and exhibits homology to the *E. coli* serine protease htrA described by  
10 Lipinska *et al.* in *Nucl. Acids Res.* 16, 10053-10066 (1988) and a putative human serine protease with an IGF-binding motif (Ohno, I., *et al.*, Genbank Accession No. D87258 (1996)), hereinafter referred to as D87258.

The *PSP1-1* sequence is listed in SEQ ID NO: 24. The coding region of this sequence consists of nucleotides 603-1979 of SEQ ID NO: 24. The deduced 458 amino  
15 acid sequence of the encoded product PSP1-1 is listed in SEQ ID NO: 25.

The *PSP1-1* sequence listed in SEQ ID NO: 30 includes two polymorphic variants, at nucleotides 672 (C/T) and 1435 (C/T) resulting in alternative amino acid residues at position 24 (arg/cys) and 278 (ala/val), both in the conserved region of nucleotides 1-1540. The deduced 458 amino acid sequence of the encoded product *PSP1-1* is listed in SEQ ID  
20 NO: 31.

The *PSP1-2* sequence is listed in SEQ ID NO: 23. The coding region of this sequence consists of nucleotides 603-1979 of SEQ ID NO: 23. The deduced 458 amino acid sequence of the encoded product PSP1-2 is listed in SEQ ID NO: 8. The *PSP1-3* sequence is listed in SEQ ID NO: 26. The coding region of this sequence consists of  
25 nucleotides 603-1736 of SEQ ID NO: 26. The deduced 377 amino acid sequence of the encoded product PSP1-3 is listed in SEQ ID NO: 27. The *PSP1-4* sequence is listed in SEQ ID NO: 28. The coding region of this sequence consists of nucleotides 603-1913 of SEQ ID NO: 28. The deduced 436 amino acid sequence of the encoded product PSP1-4 is listed in SEQ ID NO: 29.

30 The D87258 sequence is listed in SEQ ID NO: 17. The coding region of this sequence consists of nucleotides 49-1491 of SEQ ID NO: 17. The deduced 480 amino acid sequence of the encoded product D87258 is listed in SEQ ID NO: 18. The D87258 sequence listed in SEQ ID NO: 17 includes a polymorphic variant at nucleotide 1325 (G/T)



整理番号 156220

resulting in alternative amino acid residues at position 213 (gly/val). The sequence in Genbank Accession No. D87258 (1996)), describes only 1325G. The novel polynucleotide polymorph of D87258 having 1325T, is hereinafter referred to as D87258 (1325T) and the novel encoded product having valine at 213 is D87258 (1325T) protein. The novel  
 5 polynucleotide D87258 (1325T) and its encoded protein can replace *PSP-1* in any of the composition, uses or methods herein described and such novel polypeptide, encoded protein, compositions, uses and methods also form part of the invention.

As used herein, the term "functional fragments" when used to modify a specific gene or gene product means a less than full length portion of the gene or gene product  
 10 which retains substantially all of the biological function associated with the full length gene or gene product to which it relates. An example of a functional fragment of PSP1 is the minimal catalytic domain. To determine whether a fragment of a particular gene or gene product is a functional fragment, fragments are generated by well-known nucleolytic or proteolytic techniques or by the polymerase chain reaction and the fragments tested for the  
 15 described biological function.

As used herein, an "antigen" refers to a molecule containing one or more epitopes that will stimulate a host's immune system to make a humoral and/or cellular antigen-specific response. The term is also used herein interchangeably with "immunogen."

As used herein, the term "epitope" refers to the site on an antigen or hapten to  
 20 which a specific antibody molecule binds. The term is also used herein interchangeably with "antigenic determinant" or "antigenic determinant site."

As used herein, "monoclonal antibody" is understood to include antibodies derived from one species (e.g., murine, rabbit, goat, rat, human, etc.) as well as antibodies derived from two (or perhaps more) species (e.g., chimeric and humanized antibodies).

As used herein, a coding sequence is "operably linked to" another coding sequence  
 25 when RNA polymerase will transcribe the two coding sequences into a single mRNA, which is then translated into a single polypeptide having amino acids derived from both coding sequences. The coding sequences need not be contiguous to one another so long as the expressed sequence is ultimately processed to produce the desired protein.

As used herein, "recombinant" polypeptides refer to polypeptides produced by  
 30 recombinant DNA techniques; i.e., produced from cells transformed by an exogenous DNA construct encoding the desired polypeptide. "Synthetic" polypeptides are those prepared by chemical synthesis.

発明番号 156220

As used herein, a "replicon" is any genetic element (e.g., plasmid, chromosome, virus) that functions as an autonomous unit of DNA replication *in vivo*; i.e., capable of replication under its own control.

5 As used herein, a "vector" is a replicon, such as a plasmid, phage, or cosmid, to which another DNA segment may be attached so as to bring about the replication of the attached segment.

10 As used herein, a "reference" gene refers to the wild type PSP1 sequence of the invention and is understood to include the various sequence polymorphisms that exist, wherein nucleotide substitutions in the gene sequence exist, but do not affect the essential function of the gene product.

As used herein, a "mutant" gene refers to PSP1 sequences different from the reference gene wherein nucleotide substitutions and/or deletions and/or insertions result in perturbation of the essential function of the gene product.

15 As used herein, a DNA "coding sequence of" or a "nucleotide sequence encoding" a particular protein, is a DNA sequence which is transcribed and translated into a polypeptide when placed under the control of appropriate regulatory sequences.

20 As used herein, a "promoter sequence" is a DNA regulatory region capable of binding RNA polymerase in a cell and initiating transcription of a downstream (3' direction) coding sequence. For purposes of defining the present invention, the promoter sequence is bound at its 3' terminus by a translation start codon (e.g., ATG) of a coding sequence and extends upstream (5' direction) to include the minimum number of bases or elements necessary to initiate transcription at levels detectable above background. Within the promoter sequence will be found a transcription initiation site (conveniently defined by mapping with nuclease S1), as well as protein binding domains (consensus sequences) responsible for the binding of RNA polymerase. Eukaryotic promoters will often, but not always, contain "TATA" boxes and "CAT" boxes. Prokaryotic promoters contain Shine-Dalgarno sequences in addition to the -10 and -35 consensus sequences.

25 As used herein, DNA "control sequences" refers collectively to promoter sequences, ribosome binding sites, polyadenylation signals, transcription termination sequences, upstream regulatory domains, enhancers and the like, which collectively provide for the expression (i.e., the transcription and translation) of a coding sequence in a host cell.

30 As used herein, a control sequence "directs the expression" of a coding sequence in a cell when RNA polymerase will bind the promoter sequence and transcribe the coding

整理番号 156220

sequence into mRNA, which is then translated into the polypeptide encoded by the coding sequence.

As used herein, a "host cell" is a cell which has been transformed or transfected, or is capable of transformation or transfection by an exogenous DNA sequence.

5 As used herein, a cell has been "transformed" by exogenous DNA when such exogenous DNA has been introduced inside the cell membrane. Exogenous DNA may or may not be integrated (covalently linked) into chromosomal DNA making up the genome of the cell. In prokaryotes and yeasts, for example, the exogenous DNA may be maintained on an episomal element, such as a plasmid. With respect to eukaryotic cells, a stably  
10 transformed or transfected cell is one in which the exogenous DNA has become integrated into the chromosome so that it is inherited by daughter cells through chromosome replication. This stability is demonstrated by the ability of the eukaryotic cell to establish cell lines or clones comprised of a population of daughter cells containing the exogenous DNA.

15 As used herein, "transfection" or "transfected" refers to a process by which cells take up foreign DNA and integrate that foreign DNA into their chromosome. Transfection can be accomplished, for example, by various techniques in which cells take up DNA (e.g., calcium phosphate precipitation, electroporation, assimilation of liposomes, etc.) or by infection, in which viruses are used to transfer DNA into cells.

20 As used herein, a "target cell" is a cell that is selectively transfected over other cell types (or cell lines).

As used herein, a "clone" is a population of cells derived from a single cell or common ancestor by mitosis. A "cell line" is a clone of a primary cell that is capable of stable growth *in vitro* for many generations.

25 As used herein, a "heterologous" region of a DNA construct is an identifiable segment of DNA within or attached to another DNA molecule that is not found in association with the other molecule in nature. Thus, when the heterologous region encodes a gene, the gene will usually be flanked by DNA that does not flank the gene in the genome of the source animal. Another example of a heterologous coding sequence is a construct  
30 where the coding sequence itself is not found in nature (e.g., synthetic sequences having codons different from the native gene). Allelic variation or naturally occurring mutational events do not give rise to a heterologous region of DNA, as used herein.

整理番号 156220

As used herein, a "modulator" of a polypeptide is a substance which can affect the polypeptide function, such as an inhibitor of enzymatic activity.

An aspect of the present invention is isolated polynucleotides encoding a PSP1 protein and substantially similar sequences. Isolated polynucleotide sequences are substantially similar if they are capable of hybridizing under moderately stringent conditions to SEQ ID NOs: 23, 24, 26 or 28 or they encode DNA sequences which are degenerate to SEQ ID NOs: 23, 24, 26 or 28 or are degenerate to those sequences capable of hybridizing under moderately stringent conditions to SEQ ID NOs: 23, 24, 26 or 28.

Moderately stringent conditions is a term understood by the skilled artisan and has been described in, for example, Sambrook *et al. Molecular Cloning: A Laboratory Manual*, 2nd edition, Vol. 1, pp. 101-104, Cold Spring Harbor Laboratory Press (1989). An exemplary hybridization protocol using moderately stringent conditions is as follows. Nitrocellulose filters are prehybridized at 65°C in a solution containing 6X SSPE, 5X Denhardt's solution (10g Ficoll, 10g BSA and 10g polyvinylpyrrolidone per liter solution), 0.05% SDS and 100 ug/ml tRNA. Hybridization probes are labeled, preferably radiolabelled (e.g., using the Bios TAG-IT® kit). Hybridization is then carried out for approximately 18 hours at 65°C. The filters are then washed twice in a solution of 2X SSC and 0.5% SDS at room temperature for 15 minutes. Subsequently, the filters are washed at 58°C, air-dried and exposed to X-ray film overnight at -70°C with an intensifying screen.

Degenerate DNA sequences encode the same amino acid sequence as SEQ ID NOs: 8, 25, 27 or 29 or the proteins encoded by that sequence capable of hybridizing under moderately stringent conditions to SEQ ID NOs: 8, 25, 27, 29, but have variation(s) in the nucleotide coding sequences because of the degeneracy of the genetic code. For example, the degenerate codons UUC and UUU both code for the amino acid phenylalanine, whereas the four codons GGX, where X = U, C, A, or G, all code for glycine.

Alternatively, substantially similar sequences are defined as those nucleotide sequences encoding proteins having PSP1 activity in which about 70%, preferably about 80%, and most preferably about 90%, of the nucleotides share identity with *PSP1*, i.e., a sequence encoding a protein having PSP1 activity is substantially similar to any of SEQ ID NOs: 23, 24, 26 or 28 when at least about 70% of all of the nucleotides of the sequence match SEQ ID NOs: 23, 24, 26 or 28. Nucleotide sequences that are substantially similar can be identified by hybridization or by sequence comparison.

整理番号 156220

Embodiments of the isolated polynucleotides of the invention include DNA, genomic DNA and RNA, preferably of human origin. A method for isolating a nucleic acid molecule encoding a PSP1 protein is to probe a genomic or cDNA library with a natural or artificially designed probe using art recognized procedures. See, e.g., "Current Protocols in Molecular Biology", Ausubel *et al.* (eds.) Greene Publishing Association and John Wiley Interscience, New York, 1989, 1992. The ordinarily skilled artisan will appreciate that SEQ ID NOs: 23, 24, 26 or 28 or fragments thereof comprising at least 15 contiguous nucleotides are particularly useful probes. It is also appreciated that such probes can be and are preferably labeled with an analytically detectable reagent to facilitate identification of the probe. Useful reagents include, but are not limited to, radioisotopes, fluorescent dyes or enzymes capable of catalyzing the formation of a detectable product. The probes would enable the ordinarily skilled artisan to isolate complementary copies of genomic DNA, cDNA or RNA polynucleotides encoding PSP1 proteins from human, mammalian or other animal sources or to screen such sources for related sequences, e.g., additional members of the family, type and/or subtype, including transcriptional regulatory and control elements as well as other stability, processing, translation and tissue specificity-determining regions from 5' and/or 3' regions relative to the coding sequences disclosed herein, all without undue experimentation.

Another aspect of the invention is functional polypeptides encoded by the polynucleotides of the invention and substantially similar polypeptides. An embodiment of a functional polypeptide of the invention is the PSP1 protein having the amino acid sequence set forth in SEQ ID NO: 8, 25, 27 or 29.

Polypeptide sequences that are substantially similar are those sequences having PSP1 activity in which about 50%, preferably 70%, and most preferably about 90%, of the amino acids share identity with PSP1, i.e., a sequence representing a polypeptide having PSP1 activity is substantially similar to any of SEQ ID NOs: 8, 24, 26 or 28 when at least about 50% of all of the amino acids of the sequence match SEQ ID NOs: 8, 25, 27 or 29. Substantially similar polypeptide sequences can be identified by techniques such as proteolytic digestion, gel electrophoresis, microsequencing and/or sequence comparison, e.g., through use of the GAP algorithm available from the University of Wisconsin Genetics Computer Group.

Another aspect of the invention is a method for preparing essentially pure PSP1 protein. Yet another aspect is the PSP1 protein produced by the preparation method of the

整理番号 156220

invention. This protein has the amino acid sequence listed in SEQ ID NOs: 8, 25, 27 or 29 and includes variants with a substantially similar amino acid sequence that have the same function. The proteins of this invention are preferably made by recombinant genetic engineering techniques by culturing a recombinant host cell containing a vector encoding the polynucleotides of the invention under conditions promoting the expression of the protein and recovery thereof.

The isolated polynucleotides, particularly the DNAs, can be introduced into expression vectors by operatively linking the DNA to the necessary expression control regions, e.g., regulatory regions, required for gene expression. The vectors can be introduced into an appropriate host cell such as a prokaryotic, e.g., bacterial, or eukaryotic, e.g., yeast or mammalian cell by methods well known in the art. See Ausubel *et al.*, *supra*. The coding sequences for the desired proteins, having been prepared or isolated, can be cloned into any suitable vector or replicon. Numerous cloning vectors are known to those of skill in the art and the selection of an appropriate cloning vector is a matter of choice.

Examples of recombinant DNA vectors for cloning and host cells which they can transform include, but are not limited to, the bacteriophage (*E. coli*), pBR322 (*E. coli*), pACYC177 (*E. coli*), pKT230 (gram-negative bacteria), pGV1106 (gram-negative bacteria), pLAFR1 (gram-negative bacteria), pME290 (non-*E. coli* gram-negative bacteria), pHV14 (*E. coli* and *Bacillus subtilis*), pBD9 (*Bacillus*), pU61 (*Streptomyces*), pUC6 (*Streptomyces*), YIp5 (*Saccharomyces*), a baculovirus insect cell system, a *Drosophila* insect system, YCp19 (*Saccharomyces*) and pSV2neo (mammalian cells). See generally, "DNA Cloning": Vols. I & II, Glover *et al.* ed. IRL Press Oxford (1985) (1987); and T. Maniatis *et al.* ("Molecular Cloning" Cold Spring Harbor Laboratory (1982).

The gene can be placed under the control of control elements such as a promoter, ribosome binding site (for bacterial expression) and, optionally, an operator, so that the DNA sequence encoding the desired protein is transcribed into RNA in the host cell transformed by a vector containing the expression construct. The coding sequence may or may not contain a signal peptide or leader sequence. The proteins of the present invention can be expressed using, for example, the *E. coli* *tac* promoter or the protein A gene (*spa*) promoter and signal sequence. Leader sequences can be removed by the bacterial host in post-translational processing. See, e.g., U.S. Patent Nos. 4,431,739; 4,425,437 and 4,338,397.

整理番号 156220

In addition to control sequences, it may be desirable to add regulatory sequences which allow for regulation of the expression of the protein sequences relative to the growth of the host cell. Regulatory sequences are known to those of skill in the art. Exemplary are those which cause the expression of a gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound or to various temperature or metabolic conditions. Other types of regulatory elements may also be present in the vector, for example, enhancer sequences.

An expression vector is constructed so that the particular coding sequence is located in the vector with the appropriate regulatory sequences, the positioning and orientation of the coding sequence with respect to the control sequences being such that the coding sequence is transcribed under the "control" of the control sequences, i.e., RNA polymerase which binds to the DNA molecule at the control sequences transcribes the coding sequence. Modification of the sequences encoding the particular antigen of interest may be desirable to achieve this end. For example, in some cases it may be necessary to modify the sequence so that it may be attached to the control sequences with the appropriate orientation; i.e., to maintain the reading frame. The control sequences and other regulatory sequences may be ligated to the coding sequence prior to insertion into a vector, such as the cloning vectors described above. Alternatively, the coding sequence can be cloned directly into an expression vector which already contains the control sequences and an appropriate restriction site.

In some cases, it may be desirable to produce mutants or analogues of PSP1 protein. Mutants or analogues may be prepared by the deletion of a portion of the sequence encoding the protein, by insertion of a sequence, and/or by substitution of one or more nucleotides within the sequence. Techniques for modifying nucleotide sequences, such as site-directed mutagenesis, are well known to those skilled in the art. See, e.g., T. Maniatis *et al.*, *supra*; "DNA Cloning," Vols. I and II, *supra*; and "Nucleic Acid Hybridization", *supra*.

Depending on the expression system and host selected, the proteins of the present invention are produced by growing host cells transformed by an expression vector described above under conditions whereby the protein of interest is expressed. Preferred mammalian cells include human embryonic kidney cells (293), monkey kidney cells, fibroblast (COS) cells, Chinese hamster ovary (CHO) cells, *Drosophila* or murine L-cells. If the expression system secretes the protein into growth media, the protein can be purified directly from the media. If the protein is not secreted, it is isolated from cell lysates or recovered from the

整理番号 156220

cell membrane fraction. The selection of the appropriate growth conditions and recovery methods are within the skill of the art.

An alternative method to identify proteins of the present invention is by constructing gene libraries, using the resulting clones to transform *E. coli* and pooling and screening individual colonies using polyclonal serum or monoclonal antibodies to PSP1.

The proteins of the present invention may also be produced by chemical synthesis such as solid phase peptide synthesis on an automated peptide synthesizer, using known amino acid sequences or amino acid sequences derived from the DNA sequence of the genes of interest. Such methods are known to those skilled in the art.

The proteins of the present invention or their immunogenic fragments comprising at least one epitope can be used to produce antibodies, both polyclonal and monoclonal, directed to epitopes corresponding to amino acid sequences disclosed herein. If polyclonal antibodies are desired, a selected mammal such as a mouse, rabbit, goat or horse is immunized with a protein of the present invention, or its fragment, or a mutant protein. Serum from the immunized animal is collected and treated according to known procedures. Serum polyclonal antibodies can be purified by immunoaffinity chromatography or other known procedures.

Monoclonal antibodies to the proteins of the present invention, and to the immunogenic fragments thereof, can also be readily produced by one skilled in the art. The general methodology for making monoclonal antibodies by using hybridoma technology is well known. Immortal antibody-producing cell lines can be created by cell fusion and also by other techniques such as direct transformation of B lymphocytes with oncogenic DNA or transfection with Epstein-Barr virus. See, e.g., M. Schreier *et al.*, "Hybridoma Techniques" (1980); Hammerling *et al.*, "Monoclonal Antibodies and T-cell Hybridomas" (1981); Kennett *et al.*, "Monoclonal Antibodies" (1980); and U.S. Patent Nos. 4,341,761; 4,399,121; 4,427,783; 4,444,887; 4,452,570; 4,466,917; 4,472,500; 4,491,632; and 4,493,890. Panels of monoclonal antibodies produced against the antigen of interest, or fragment thereof, can be screened for various properties, i.e., for isotype, epitope, affinity, etc. Monoclonal antibodies are useful in purification, using immunoaffinity techniques, of the individual antigens which they are directed against. Alternatively, genes encoding the monoclonals of interest may be isolated from the hybridomas by PCR techniques known in the art and cloned and expressed in the appropriate vectors. The antibodies of this invention, whether polyclonal or monoclonal have additional utility in that they may be employed as reagents



整理番号 156220

in immunoassays, RIA, ELISA, and the like. The antibodies of the invention can be labeled with an analytically detectable reagent such as a radioisotope, fluorescent molecule or enzyme.

Chimeric antibodies, in which non-human variable regions are joined or fused to human constant regions (see, e.g., Liu *et al.*, *Proc. Natl Acad. Sci. USA*, 84, 3439 (1987)), may also be used in assays or therapeutically. Preferably, a therapeutic monoclonal antibody would be "humanized" as described in Jones *et al.*, *Nature*, 321, 522 (1986); Verhoeven *et al.*, *Science*, 239, 1534 (1988); Kabat *et al.*, *J. Immunol.*, 147, 1709 (1991); Queen *et al.*, *Proc. Natl Acad. Sci. USA*, 86, 10029 (1989); Gorman *et al.*, *Proc. Natl Acad. Sci. USA*, 88, 34181 (1991); and Hodgson *et al.*, *BioTechnology*, 9, 421 (1991).

Another aspect of the present invention is modulators of the polypeptides of the invention or of D87258. Functional modulation of PSP1 or D87258 by a substance includes partial to complete inhibition of function, such as inhibition of proteolytic activity, identical function, as well as enhancement of function. Embodiments of modulators of the invention include peptides, oligonucleotides and small organic molecules including peptidomimetics. Modulators of the invention may be useful as therapeutics or prophylactics for all forms of neurodegeneration including AD. Modulators of PSP1 or D87258 proteolytic activity relative to other endogenous substrates may be also be useful for the treatment of other types of human disease states.

Another aspect of the invention is antisense oligonucleotides comprising a sequence which is capable of binding to the polynucleotides of the invention. Synthetic oligonucleotides or related antisense chemical structural analogs can be designed to recognize, specifically bind to and prevent transcription of a target nucleic acid encoding PSP1 or D87258 protein by those of ordinary skill in the art. See generally, Cohen, J.S., *Trends in Pharm. Sci.*, 10, 435(1989) and Weintraub, H.M., *Scientific American*, January (1990) at page 40.

Another aspect of the invention is a method for assaying a medium for the presence of a substance that modulates PSP1 or D87258 protein function by affecting the binding of PSP1 or D87258 protein to cellular binding partners. Examples of modulators include, but are not limited to peptides and small organic molecules including peptidomimetics. A PSP1 or D87258 protein is provided having the amino acid sequence of PSP1 (SEQ ID NOs: 8, 25, 27 or 29) or D87258 (SEQ ID NO: 18) or a functional fragment thereof together with a cellular binding partner or synthetic analog thereof. The mixture is

整理番号 156220

incubated with a test substance which is suspected of modulating PSP1 or D87258 activity, under conditions which permit the formation of a PSP1 or D87258 gene product/cellular binding partner complex. An assay is performed for the presence of the complex, free PSP1 or D87258 protein or free cellular binding partner and the result compared to a control to determine the effect of the test substance.

Another aspect of the invention is a method for assaying a medium for the presence of a substance that modulates PSP1 or D87258 protein function by inhibiting its proteolytic activity on cellular substrates. Examples of modulators include, but are not limited to peptides and small organic molecules including peptidomimetics. Cellular substrates can include PS-1, PS-2, APP or other substrates. A PSP1 or D87258 protein is provided having the amino acid sequence of PSP1 (SEQ ID NOs: 8, 25, 27 or 29) or D87258 (SEQ ID NO: 18) or a functional fragment thereof together with a cellular substrate or synthetic analog thereof. The mixture is incubated with a test substance which is suspected of inhibiting PSP1 or D87258 activity, under conditions which permit the formation of a PSP1 or D87258 enzyme/substrate complex and subsequent cleavage of the substrate.

整型番号 156220

Another aspect of the invention is a method for assaying for the presence of a substance that modulates PSP1 or D87258 activity by direct binding to PSP1 or D87258 protein. Examples of modulators include, but are not limited to, peptides and small organic molecules including peptidomimetics. Modulator candidates are synthesized on a solid support by techniques such as those disclosed in Lam *et al.*, *Nature* 354, 82 (1991) or 5 Burbaum *et al.*, *Proc. Natl. Acad. Sci. USA* 92, 6027 (1995) to provide solid support-associated modulator candidates. A labelled PSP1 or D87258 protein is provided having the amino acid sequence of PSP1 (SEQ ID NOs: 8, 25, 27 or 29) or D87258 (SEQ ID NO: 18) or a functional derivative thereof. Exemplary labels include directly attached 10 fluorescent or colored dyes, biotin, radioisotopes or epitope tags, which are detectable by a suitable antibody. A mixture of solid support-associated modulator candidates and labelled PSP1 or D87258 protein is incubated under conditions which can permit the formation of a PSP1 or D87258 protein/modulator candidate complex. The solid support is separated from free soluble labelled PSP1 or D87258 protein. An assay is performed for the presence of 15 solid support-associated labelled protein. Solid supports complexed with labelled protein are isolated and the identity of the modulator candidate determined by techniques well known to those skilled in the art, such as the TOF-SIMS method in Brummel *et al.*, *Science* 264, 399-402 (1994).

Modulation of PSP1 or D87258 function would be expected to have effects on 20 presenilin cleavage, the cleavage of other proteins or  $\beta$ A4 production. Any modulators so identified would be expected to be useful as a therapeutic for the treatment and prevention of neurodegeneration including FAD and AD.

Further, PSP1 or D87258 could be used to isolate proteins which interact with it and this interaction could be a target for interference. Inhibitors of protein-protein 25 interactions between PSP1 or D87258 and other factors could lead to the development of pharmaceutical agents for the modulation of PSP1 or D87258 activity.

Methods to assay for protein-protein interactions, such as that of a PSP1 or D87258 gene product/binding partner complex, and to isolate proteins interacting with PSP1 or D87258 are known to those skilled in the art. Use of the methods discussed below enable 30 one of ordinary skill in the art to accomplish these aims without undue experimentation.

The yeast two-hybrid system provides methods for detecting the interaction between a first test protein and a second test protein, *in vivo*, using reconstitution of the activity of a transcriptional activator. The method is disclosed in U.S. Patent No.

整理番号 156220

5,283,173; reagents are available from Clontech and Stratagene. Briefly, PSP1 cDNA is fused to a *Gal4* or *LexA* transcription factor DNA binding domain and expressed in yeast cells. cDNA library members obtained from cells of interest are fused to a transactivation domain of *Gal4* or another transactivation domain. cDNA clones which express proteins which can interact with PSP1 will lead to reconstitution of transcription factor activity such as *Gal4* and transactivation of a reporter gene expression such as *Gal1-lacZ*.

10 An alternative method is screening of  $\lambda$ gt11,  $\lambda$ ZAP (Stratagene) or equivalent cDNA expression libraries with recombinant PSP1. Recombinant PSP1 protein or fragments thereof are fused to small peptide tags such as FLAG, HSV or GST. The peptide tags can possess convenient phosphorylation sites for a kinase such as heart muscle creatine kinase or they can be biotinylated. Recombinant PSP1 can be phosphorylated with  $^{32}$ [P] or used unlabeled and detected with streptavidin or antibodies against the tags.  $\lambda$ gt11 cDNA expression libraries are made from cells of interest and are incubated with the recombinant PSP1, washed and cDNA clones isolated which interact with PSP1. See, e.g., T. Maniatis  
15 *et al.*, *supra*.

Another method is the screening of a mammalian expression library in which the cDNAs are cloned into a vector between a mammalian promoter and polyadenylation site and transiently transfected in COS or 293 cells followed by detection of the binding protein 48 hours later by incubation of fixed and washed cells with a labelled PSP1, preferably iodinated, and detection of bound PSP1 by autoradiography (See Sims *et al.*, *Science* 241, 585-589 (1988) and McMahan *et al.*, *EMBO J.* 10, 2821-2832 (1991)). In this manner, pools of cDNAs containing the cDNA encoding the binding protein of interest can be selected and the cDNA of interest can be isolated by further subdivision of each pool followed by cycles of transient transfection, binding and autoradiography. Alternatively,  
25 the cDNA of interest can be isolated by transfecting the entire cDNA library into mammalian cells and panning the cells on a dish containing PSP1 bound to the plate. Cells which attach after washing are lysed and the plasmid DNA isolated, amplified in bacteria, and the cycle of transfection and panning repeated until a single cDNA clone is obtained (See Seed *et al.*, *Proc. Natl. Acad. Sci. USA* 84, 3365 (1987) and Aruffo *et al.*, *EMBO J.* 6, 3313 (1987)). If the binding protein is secreted, its cDNA can be obtained by a similar  
30 pooling strategy once a binding or neutralizing assay has been established for assaying supernatants from transiently transfected cells. General methods for screening supernatants are disclosed in Wong *et al.*, *Science* 228, 810-815 (1985).

整理番号 156220

Another alternative method is isolation of proteins interacting with PSP1 directly from cells. Fusion proteins of PSP1 with GST or small peptide tags are made and immobilized on beads. Biosynthetically labeled or unlabeled protein extracts from the cells of interest are prepared, incubated with the beads and washed with buffer. Proteins  
 5 interacting with PSP1 are eluted specifically from the beads and analyzed by SDS-PAGE. Binding partner primary amino acid sequence data are obtained by microsequencing. Optionally, the cells can be treated with agents that induce a functional response such as tyrosine phosphorylation of cellular proteins. An example of such an agent would be a growth factor or cytokine such as erythropoietin or Interleukin-3.

10 Another alternative method is immunoaffinity purification. Recombinant PSP1 is incubated with labeled or unlabeled cell extracts and immunoprecipitated with anti-PSP1 antibodies. The immunoprecipitate is recovered with protein A-Sepharose and analyzed by SDS-PAGE. Unlabelled proteins are labeled by biotinylation and detected on SDS gels with streptavidin. Binding partner proteins are analyzed by microsequencing. Further,  
 15 standard biochemical purification steps known to those skilled in the art may be used prior to microsequencing.

Yet another alternative method is screening of peptide libraries for binding partners. Recombinant tagged or labeled PSP1 is used to select peptides from a peptide or phosphopeptide library which interact with PSP1. Sequencing of the peptides leads to  
 20 identification of consensus peptide sequences which might be found in interacting proteins.

PSP1 or D87258 binding partners identified by any of these methods or other methods which would be known to those of ordinary skill in the art as well as those putative binding partners discussed above can be used in the assay method of the invention. Assaying for the presence of PSP1 or D87258 /binding partner complex are accomplished  
 25 by, for example, the yeast two-hybrid system, ELISA or immunoassays using antibodies specific for the complex. In the presence of test substances which interrupt or inhibit formation of PSP1 or D87258 /binding partner interaction, a decreased amount of complex will be determined relative to a control lacking the test substance.

Assays for free PSP1 or D87258, or binding partner are accomplished by, for  
 30 example, ELISA or immunoassay using specific antibodies or by incubation of radiolabeled PSP1 or D87258 with cells or cell membranes followed by centrifugation or filter separation steps. In the presence of test substances which interrupt or inhibit formation of

整理番号 156220

PSP1 or D87258 /binding partner interaction, an increased amount of free PSP1 or D87258, or free binding partner will be determined relative to a control lacking the test substance.

Another aspect of the invention is pharmaceutical compositions comprising an effective amount of a PSP1 or D87258 modulator of the invention and a pharmaceutically acceptable carrier. Pharmaceutical compositions of modulators of this invention for parenteral administration, i.e., subcutaneously, intramuscularly or intravenously or oral administration can be prepared.

The compositions for parenteral administration will commonly comprise a solution of the modulators of the invention or a cocktail thereof dissolved in an acceptable carrier, preferably an aqueous carrier. A variety of aqueous carriers may be employed, e.g., water, buffered water, 0.4% saline, 0.3% glycine and the like. These solutions are sterile and generally free of particulate matter. These solutions may be sterilized by conventional, well-known sterilization techniques. The compositions may contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions such as pH adjusting and buffering agents, etc. The concentration of the modulator of the invention in such pharmaceutical formulation can vary widely, i.e., from less than about 0.5%, usually at or at least about 1% to as much as 15 or 20% by weight and will be selected primarily based on fluid volumes, viscosities, etc. according to the particular mode of administration selected.

Thus, a pharmaceutical composition of the modulator of the invention for intramuscular injection could be prepared to contain 1 mL sterile buffered water, and 50 mg of a protein of the invention. Similarly, a pharmaceutical composition of the modulator of the invention for intravenous infusion could be made up to contain 250 mL of sterile Ringer's solution, and 150 mg of a modulator of the invention. Actual methods for preparing parenterally administrable compositions are well known or will be apparent to those skilled in the art and are described in more detail in, for example, *Remington's Pharmaceutical Science*, 15th ed., Mack Publishing Company, Easton, Pennsylvania.

The physician will determine the dosage of the present therapeutic agents which will be most suitable and it will vary with the form of administration and the particular compound chosen, and furthermore, it will vary with the particular patient under treatment. Generally, the physician will wish to initiate treatment with small dosages substantially less than the optimum dose of the compound and increase the dosage by small increments until the optimum effect under the circumstances is reached. It will generally be found that

整理番号 156220

when the composition is administered orally, larger quantities of the active agent will be required to produce the same effect as a smaller quantity given parenterally. The therapeutic dosage will generally be from 0.1 to 1000 milligrams per day and higher although it may be administered in several different dosage units.

- 5            Depending on the patient condition, the pharmaceutical composition of the invention can be administered for prophylactic and/or therapeutic treatments. In therapeutic applications, compositions containing the present compounds or a cocktail thereof are administered to a patient already suffering from a disease in an amount sufficient to cure or at least partially arrest the disease and its complications. In
- 10 prophylactic applications, compositions containing the present compounds or a cocktail thereof are administered to a patient not already in a disease state to enhance the patient's resistance to the disease.

- Single or multiple administrations of the pharmaceutical compositions can be carried out with dose levels and pattern being selected by the treating physician. In any
- 15 event, the pharmaceutical composition of the invention should provide a quantity of the modulators of the invention sufficient to effectively treat the patient.

- Additionally, some diseases result from inherited defective genes. These genes can be detected by comparing the sequence of the defective gene with that of a normal one. Individuals carrying mutations in the PSP1 or D87258 gene may be detected at the DNA
- 20 level by a variety of techniques. Nucleic acids used for diagnosis (genomic DNA, mRNA, etc.) may be obtained from a patient's cells, such as from blood, urine, saliva or tissue biopsy, e.g., chorionic villi sampling or removal of amniotic fluid cells and autopsy material. The genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR, ligase chain reaction (LCR), strand displacement
- 25 amplification (SDA), etc. prior to analysis. See, e.g., Saiki *et al.*, *Nature*, 324, 163-166 (1986), Bej, *et al.*, *Crit. Rev. Biochem. Molec. Biol.*, 26, 301-334 (1991), Birkenmeyer *et al.*, *J. Virol. Meth.*, 35, 117-126 (1991), Van Brunt, J., *BioTechnology*, 8, 291-294 (1990)). RNA or cDNA may also be used for the same purpose. As an example, PCR primers complementary to the nucleic acid of the instant invention can be used to identify and
- 30 analyze PSP1 or D87258 mutations. For example, deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal PSP1 or D87258 genotype. Point mutations can be identified by hybridizing amplified DNA to radiolabeled PSP1 or D87258 RNA of the invention or alternatively, radiolabelled PSP1 or D87258

整理番号 156220

antisense DNA sequences of the invention. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase A digestion or by differences in melting temperatures ( $T_m$ ). Such a diagnostic would be particularly useful for prenatal and even neonatal testing.

- 5           In addition, point mutations and other sequence differences between the reference gene and "mutant" genes can be identified by yet other well-known techniques, e.g., direct DNA sequencing, single-strand conformational polymorphism. See Orita *et al.*, *Genomics*, 5, 874-879 (1989). For example, a sequencing primer is used with double-stranded PCR product or a single-stranded template molecule generated by a modified PCR. The
- 10           sequence determination is performed by conventional procedures with radiolabeled nucleotides or by automatic sequencing procedures with fluorescent-tags. Cloned DNA segments may also be used as probes to detect specific DNA segments. The sensitivity of this method is greatly enhanced when combined with PCR. Further, point mutations and other sequence variations, such as polymorphisms, can be detected as described above, e.g.,
- 15           through the use of allele-specific oligonucleotides for PCR amplification of sequences that differ by single nucleotides. Oligonucleotides having sequences as set forth in SEQ ID Nos: 32, 33, 34, 35, 36, 37, 38, 39 and 40 are useful in such a method. These methods are useful for determining the genetic predisposition to neurodegeneration in a patient by detecting polymorphisms within PSP1 or D87258 in a sample from a patient. Preferably,
- 20           the polymorphisms detected are at nucleotide 672 of PSP1, at nucleotide 1435 of PSP1 or at nucleotide 1325 of D87258. Preferably, the polymorphisms are detected by PCR; most preferably, the polymorphisms are detected by PCR with oligonucleotides having a nucleotide sequence selected from the group consisting of SEQ ID NOs: 32, 33, 34, 35, 36, 37, 38, 39 and 40. Preferably, the neurodegeneration predisposition determined is to
- 25           Alzheimer's disease.

- Genetic testing based on DNA sequence differences may be achieved by detection of alteration in electrophoretic mobility of DNA fragments in gels with or without denaturing agents. Small sequence deletions and insertions can be visualized by high resolution gel electrophoresis. DNA fragments of different sequences may be distinguished
- 30           on denaturing formamide gradient gels in which the mobilities of different DNA fragments are retarded in the gel at different positions according to their specific melting or partial melting temperatures. See, e.g., Myers *et al.*, *Science*, 230, 1242 (1985). In addition, sequence alterations, in particular small deletions, may be detected as changes in the



整理番号 156220

migration pattern of DNA heteroduplexes in non-denaturing gel electrophoresis such as heteroduplex electrophoresis. See, e.g., Nagamine *et al.*, *Am. J. Hum. Genet.*, 45, 337-339 (1989). Sequence changes at specific locations may also be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method as disclosed by  
 5 Cotton *et al.* in *Proc. Natl. Acad. Sci. USA*, 85, 4397-4401 (1985).

Thus, the detection of a specific DNA sequence may be achieved by methods such as hybridization (e.g., heteroduplex electroporation, see, White *et al.*, *Genomics*, 12, 301-306 (1992), RNase protection (e.g., Myers *et al.*, *Science*, 230, 1242 (1985)) chemical cleavage (e.g., Cotton *et al.*, *Proc. Natl. Acad. Sci. USA*, 85, 4397-4401 (1985)), direct  
 10 DNA sequencing, or the use of restriction enzymes (e.g., restriction fragment length polymorphisms (RFLP) in which variations in the number and size of restriction fragments can indicate insertions, deletions, presence of nucleotide repeats and any other mutation which creates or destroys an endonuclease restriction sequence). Southern blotting of genomic DNA may also be used to identify large (i.e., greater than 100 base pair) deletions  
 15 and insertions.

In addition to conventional gel electrophoresis and DNA sequencing, mutations such as microdeletions, aneuploidies, translocations, inversions, can also be detected by *in situ* analysis. See, e.g., Keller *et al.*, *DNA Probes*, 2nd Ed., Stockton Press, New York, N.Y., USA (1993). That is, DNA or RNA sequences in cells can be analyzed for mutations  
 20 without isolation and/or immobilization onto a membrane. Fluorescence *in situ* hybridization (FISH) is presently the most commonly applied method and numerous reviews of FISH have appeared. See, e.g., Trachuck *et al.*, *Science*, 250, 559-562 (1990), and Trask *et al.*, *Trends, Genet.*, 7, 149-154 (1991). Hence, by using nucleic acids based on the structure of the PSP1 or D87258 genes, one can develop diagnostic tests for genetic  
 25 mutations.

In addition, some diseases are a result of, or are characterized by, changes in gene expression which can be detected by changes in the mRNA. Alternatively, the PSP1 or D87258 gene can be used as a reference to identify individuals expressing an increased or decreased level of PSP1 or D87258 mRNA, e.g., by Northern blotting or *in situ*  
 30 hybridization.

Defining appropriate hybridization conditions is within the skill of the art. See, e.g., "Current Protocols in Mol. Biol." Vol. I & II, Wiley Interscience. Ausbel *et al.* (eds.) (1992). Probing technology is well known in the art and it is appreciated that the size of the

整理番号 156220

probes can vary widely but it is preferred that the probe be at least 15 nucleotides in length. It is also appreciated that such probes can be and are preferably labeled with an analytically detectable reagent to facilitate identification of the probe. Useful reagents include but are not limited to radioisotopes, fluorescent dyes or enzymes capable of catalyzing the  
 5 formation of a detectable product. As a general rule, the more stringent the hybridization conditions the more closely related genes will be that are recovered.

The putative role of PSP1 or D87258 in presenilin biochemistry establishes yet another aspect of the invention which is gene therapy. "Gene therapy" means gene supplementation where an additional reference copy of a gene of interest is inserted into  
 10 patient's cells. As a result, the protein encoded by the reference gene corrects the defect and permits the cells to function normally, thus alleviating disease symptoms. The reference copy would be a wild-type form of the PSP1 or D87258 gene or a gene encoding a protein or peptide which modulates the activity of the endogenous PSP1 or D87258.

Gene therapy of the present invention can occur *in vivo* or *ex vivo*. *Ex vivo* gene  
 15 therapy requires the isolation and purification of patient cells, the introduction of a therapeutic gene and introduction of the genetically altered cells back into the patient. A replication-deficient virus such as a modified retrovirus can be used to introduce the therapeutic PSP1 or D87258 gene into such cells. For example, mouse Moloney leukemia virus (MMLV) is a well-known vector in clinical gene therapy trials. See, e.g., Boris-  
 20 Lauerie *et al.*, *Curr. Opin. Genet. Dev.*, 3, 102-109 (1993).

In contrast, *in vivo* gene therapy does not require isolation and purification of a patient's cells. The therapeutic gene is typically "packaged" for administration to a patient such as in liposomes or in a replication-deficient virus such as adenovirus as described by Berkner, K.L., in *Curr. Top. Microbiol. Immunol.*, 158, 39-66 (1992) or adeno-associated  
 25 virus (AAV) vectors as described by Muzyczka, N., in *Curr. Top. Microbiol. Immunol.*, 158, 97-129 (1992) and U.S. Patent No. 5,252,479. Another approach is administration of "naked DNA" in which the therapeutic gene is directly injected into the bloodstream or muscle tissue. Another approach is administration of "naked DNA" in which the therapeutic gene is introduced into the target tissue by microparticle bombardment using  
 30 gold particles coated with the DNA.

Cell types useful for gene therapy of the present invention include lymphocytes, hepatocytes, myoblasts, fibroblasts, any cell of the eye such as retinal cells, epithelial and endothelial cells. Preferably the cells are T lymphocytes drawn from the patient to be

整理番号 156220

treated, hepatocytes, any cell of the eye or respiratory or pulmonary epithelial cells. Transfection of pulmonary epithelial cells can occur via inhalation of a nebulized preparation of DNA vectors in liposomes, DNA-protein complexes or replication-deficient adenoviruses. See, e.g., U.S. Patent No. 5,240,846.

5 Another aspect of the invention is transgenic, non-human mammals capable of expressing the polynucleotides of the invention or D87258 in any cell. Transgenic, non-human animals may be obtained by transfecting appropriate fertilized eggs or embryos of a host with the polynucleotides of the invention, with D87258 or with mutant forms found in human diseases. See, e.g., U.S. Patent Nos. 4,736,866; 5,175,385; 5,175,384 and  
10 5,175,386. The resultant transgenic animal may be used as a model for the study of *PSP1* or *D87258* gene function. Particularly useful transgenic animals are those which display a detectable phenotype associated with the expression of the *PSP1* or *D87258* protein. Drug development candidates may then be screened for their ability to reverse or exacerbate the relevant phenotype.

15

The present invention will now be described with reference to the following specific, non-limiting examples.

#### Example 1 - Identification of the PS-1 Binding Partner PSP1

20 A portion of PS-1 cDNA (GenBank Accession No. L42110) (SEQ ID NO: 9) encoding residues 269-413 of the PS-1 amino acid sequence (SEQ ID NO: 10) was PCR amplified with the oligonucleotide primers 5'-CGGAATTCGATGCTGGTTGAAACA-3' (SEQ ID NO: 11) and 5'-CGGGATCCTCAGGCTACGAAACAGGCTAT-3' (SEQ ID NO: 12). The product was digested with *EcoRI* and *BamHI* and cloned into pEG202  
25 (Golemis *et al.*, in *Current Protocols in Molecular Biology*, John Wiley & Sons, New York (1994)). The resulting plasmid, pCC352, encoded a fusion protein in which the DNA binding protein, *LexA*, was fused in-frame to amino acids 269-413 of PS-1. The parent vector, pEG202, was a yeast expression vector which uses the alcohol dehydrogenase (*ADHI*) promoter to express the *LexA* fusion proteins and *HIS3* as the selectable marker.  
30 Sequence analysis using an automated DNA sequencer (Applied Biosystems, Inc.) confirmed that the amplified region had the correct sequence and was fused in-frame to *LexA*.

整理番号 156220

All procedures, plasmids and strains used in the two-hybrid screen have been described in detail by Golemis *et al.*, *supra*. Yeast strain EGY48 (*MATa*, *trp1*, *his3*, *ura3*, 6ops-*LEU2*) was cotransformed with the plasmids pCC352 and pSH18-34. Transformants were selected using complete minimal media lacking uracil and histidine. The plasmid pSH18-34 is a yeast expression vector in which eight LexA operator sites are located upstream of a minimal *GALI* promoter which drives the expression of the *LacZ* gene and *URA3* as a selectable marker. Synthesis of the full length LexA-PS-1 fusion was confirmed by Western blot analysis of yeast extracts using polyclonal antisera directed against LexA. It was confirmed that the LexA-PS-1 fusion alone was unable to activate neither the *LEU2* nor *LacZ* reporter strains. In addition, the ability of the LexA-PS-1 fusion to enter the nucleus and bind DNA was confirmed using a repression assay.

A strain containing the LexA-PS-1 fusion and pSH18-34 (CCY321) was transformed with a human fetal brain cDNA library (Clontech) in plasmid pJG4-5 using a library scale transformation protocol. This library plasmid contains the *TRP1* selectable marker and allows the expression of cDNAs as fusions (AD fusions) to a cassette containing the SV40 nuclear localization sequence, the acid blob B42, and the hemagglutinin epitope tag. See Gyuris *et al.*, *Cell* 75, 791-803 (1993). Expression of this fusion is under control of the galactose inducible promoter *GALI*. Transformation reactions were plated onto complete minimal media lacking uracil, histidine and tryptophan. Approximately  $4.5 \times 10^6$  individual transformants were obtained, pooled and frozen. To ensure that each primary colony was replated during the selection procedure,  $2 \times 10^7$  viable cells (approximately 3 times the number of individual transformants) were plated onto minimal media lacking uracil, histidine, tryptophan and leucine with galactose/raffinose as the carbon source to induce expression of AD fusions. Colonies arising after 3 and 4 days of growth at 30 °C were picked to complete minimal media lacking uracil, histidine and tryptophan. Colonies containing potential interacting fusion proteins were then tested for galactose dependence and LacZ expression. Those isolates which activated both the *LEU2* and *LacZ* reporters in a galactose dependent fashion were considered positive and pursued further. Plasmids were isolated from yeast, used to transform *E. coli* strain KC8, and AD fusion plasmids selected by growth on minimal *E. coli* media lacking tryptophan. Each AD fusion plasmid containing a potential interacting fusion was used to transform CCY321. Several transformants were subjected to screening for galactose dependent *LEU2* and *LacZ* activation. To ensure that the interaction was

登録番号 156220

specific, the ability of each AD fusion plasmid to interact with 22 nonrelated LexA fusion proteins was tested. AD fusion plasmids which passed this second round of screening and interacted specifically with the LexA-PS-1 fusion were identified.

## 5 Example 2 - PSP1 cDNA Cloning and Sequence Analysis

The AD fusion plasmids were subjected to restriction digest analysis and sequencing as indicated above. Sequence analysis of one of the interacting fusion protein cDNAs revealed a 519 nucleotide open reading frame (SEQ ID NO: 1) encoding a 173 amino acid (SEQ ID NO: 2) protein starting with an GGA at position 2 and terminating with a TGA at position 523 of SEQ ID NO: 1. GenBank searches using the BLASTX and BLASTN algorithms with the cDNA sequence or with the deduced amino acid sequence indicated homology to a portion of the *E. coli* serine protease htrA described by Lipinska *et al.*, *supra*, (SEQ ID NOs: 13 and 14). This novel cDNA was designated PSP1.

To obtain a greater portion of the cDNA, the oligonucleotide, 5'-CTGGATGGGGAGGTGATTGGAGTG-3' (SEQ ID NO: 15) representing bp 83-106 of SEQ ID NO: 1, was used to screen a Superscript human brain cDNA library (Gibco BRL) using the Genetrappor cDNA positive selection system (Gibco BRL). Colonies were screened using whole cell PCR or standard hybridization conditions as described by Innis *et al.*, *PCR Protocols: A Guide to Methods and Applications*, Academic Press, San Diego, CA (1990) and Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Press, Cold Spring Harbor, NY (1989). Those isolates which contained PSP1 were subjected to restriction digest analysis and sequencing. The longest clones, SEQ ID NO: 3 and SEQ ID NO: 5 were sequenced in their entirety.

Sequence analysis of SEQ ID NO: 3 revealed a 969 nucleotide open reading frame encoding a 323 amino acid (SEQ ID NO: 4) protein starting with a CCC at position 1 and terminating with a TGA at position 972 of SEQ ID NO: 3. Sequence analysis of SEQ ID NO: 5 revealed a 1500 nucleotide open reading frame encoding a 423 amino acid (SEQ ID NO: 6) protein starting with an CTT at position 1 and terminating with a TGA at position 1272 of SEQ ID NO: 5.

A second round of screening was performed using the oligonucleotide, 5'-GTCTCTGGGCCCCGGTTGTCTGTTG-3' (SEQ ID NO: 16) representing bp 5-28 of SEQ ID NO: 5; the library and screening protocol remained unchanged. In the second round of screening, the isolate designated SEQ ID NO: 7 contained the longest cDNA clone.

整理番号 156220

Sequence analysis of SEQ ID NO: 7 revealed a 1374 nucleotide open reading frame encoding a 458 amino acid (SEQ ID NO: 8) protein starting with an ATG at position 251 and terminating with a TGA at position 1627 of SEQ ID NO: 7. However, SEQ ID NO: 7 does not have a stop codon upstream from the potential initiation codon. To confirm that the predicted start codon is authentic, the 5' nucleotide sequence was extended with 5' RACE using "Marathon Ready" human brain cDNA (Clontech) and a nested set of primers. A SEQ ID NO: 7 specific primer 5'-CCAACAGACAACCGGGCCCAGAGACT-3' (SEQ ID NO: 20) and a 5' anchor primer-1 (Clontech) was used in the first PCR amplification and a SEQ ID NO: 7 specific primer 5'-TGCCTCCTCGCCCGCCCTACTCAGA-3' (SEQ ID NO: 21) and 5' anchor primer-2 (Clontech) was used in the second PCR amplification. PCR products were T/A cloned into pCR2.1 (Invitrogen). Eighteen isolates with staggered 5' ends were analyzed and a 5' consensus sequence of 587 nucleotides was generated (SEQ ID NO: 22). Alignment of SEQ ID NO: 22 and SEQ ID NO: 7 to generate a consensus sequence (SEQ ID NO: 23) indicates that at nucleotide position 225 there is an in frame stop codon and the first methionine corresponds to that predicted in SEQ ID NO: 7. This gene is designated *PSP1-2*.

Consensus full length sequences for the genes designated *PSP1-1* (SEQ ID NOs: 24 and 25), *PSP1-3* (SEQ ID NOs: 26 and 27) and *PSP1-4* (SEQ ID NOs: 28 and 29) were generated from alignment of the 5' consensus sequence (SEQ ID NO: 22), other partial *PSP1* clones, and with SEQ ID NOs: 7, 3 and 5, respectively.

Alignment of the deduced amino acid sequence of *PSP1-1* (SEQ ID NO: 25) to *E. coli* *htrA* (SEQ ID NO: 14) was accomplished using the BESTFIT algorithm (University of Wisconsin Genetics Computer Group). An approximate similarity of 55% and an identity of 33.5% at the amino acid level was observed and is shown in Fig. 1 (top, *PSP1-1*; bottom, *E. coli* *htrA*). The critical histidine and serine motif GX SXG conserved in all serine proteases is present in *PSP1-1* at amino acid positions 198 and 304-308, respectively, and are indicated in bold. Amino acid numbers are indicated at the left and right of the sequence alignment.

Nucleotide sequence comparison of *PSP1-2*, *PSP1-1*, *PSP1-3* and *PSP1-4* using the PILEUP and PRETTY algorithms (University of Wisconsin Genetics Computer Group) with gap creation and extension penalties of 5.0 and 0.3, respectively, is shown in Fig. 2. The alignment results indicate that at nucleotide position 1541 of the alignment, *PSP1-2* and *PSP1-1* contain a 225 bp deletion and *PSP1-4* contains a 195 bp deletion. Within the

整理番号 156220

- same alignment at nucleotide position 1942, *PSP1-4* lacks 96 bp that are present in *PSP1-2*, *PSP1-1* and *PSP1-3*. At the junction of each deletion site there is a splice site consensus sequence AGG or TGG (indicated in bold), suggesting that these alternate forms are due to alternative splicing. See Mount, S. in *Nucl. Acids Res* 10, 458-472 (1982). The apparent
- 5 splicing event at position 1541 results in the removal of a stop codon (underlined in Fig. 2) that is present in *PSP1-3*. In addition, *PSP1-2* and *PSP1-1* contain a single nucleotide difference at position 672 of the alignment. *PSP1-2* contains a T at this position producing the codon TGC which codes for a cysteine while *PSP1-1* contains a C at the same position producing the codon CGC which codes for a cysteine.
- 10 Nucleotide sequence comparison of *PSP1-1* (SEQ ID NO: 24) to the putative human serine protease of Ohno *et al.*, *supra*, (SEQ ID NO: 17) indicated a 49% identity using the GAP algorithm and 65% using the BESTFIT algorithm (data not shown). Alignment of the deduced amino acid sequence of *PSP1-1* (SEQ ID NO: 25) to the D87258
- 15 protease of Ohno *et al.*, *supra*, (SEQ ID NO: 18) was accomplished using the BESTFIT algorithm and is shown in Fig. 3 (top, *PSP1-1*; bottom, Ohno *et al.* D87258 protease). An approximate identity of 46% at the amino acid level was observed.

### Example 3 - Tissue Distribution of PSP1

- Northern analysis was carried out to determine the distribution of *PSP1* mRNA in
- 20 human tissues. A 30-base oligonucleotide probe directed against the *PSP1* sequence was used (5'-ATGCTGAACATCGGGAAAGCTTGTTCTCG-3') (SEQ ID NO: 19). This probe was 3'-end labelled with [<sup>32</sup>P]-dATP. Northern blots containing mRNA from multiple human tissues (Clontech #7750-1, #7760-1, and #7755-1) were hybridized with this probe under stringent conditions. A major band of approximately 1.9kb was detected
- 25 in all regions investigated: heart, brain, lung, placenta, liver, skeletal muscle, kidney, pancreas, amygdala, caudate nucleus, corpus callosum, hippocampus, substantia nigra, subthalamic nucleus, thalamus, cerebellum, cerebral cortex, medulla, spinal cord, occipital pole, frontal lobe, temporal pole, and putamen. *PSP1* mRNA was also detected in Alzheimer's disease brain.

30

登録番号 156220

#### Example 4 - Detecting the PSP1 polymorphisms

PSP1 oligonucleotides 1AFC, 1AFT and 1AR were designed for detecting the polymorphism at nucleotide 672 (cytidine to thymine) causing the Arg to Cys amino acid change. The Allele Specific Oligonucleotides (ASO) 1AFC and 1AFT are identical apart  
5 from their 3' end bases and provide the specificity for screening for the polymorphism.

1AFC: CAT CCG GCA TTG TTA GCT CTG C 22mer (SEQ ID NO:32)

1AFT: CAT CCG GCA TTG TTA GCT CTG T 22mer (SEQ ID NO:33)

1AR: CAA TAG CTG CAT CAG TTT GAA TG 23mer (SEQ ID NO:34)

10 Pairs of oligonucleotides (1AFC + 1AR, or 1AFT + 1AR) were used in a PCR under the following conditions: 94°C for 40 seconds, 60°C for 30 seconds, for 35 cycles in a reaction containing 1 U KlenTaq1 (GenPak Ltd.), 50mM Tris-Cl pH9.1, 16mM ammonium sulphate, 3.5mM MgCl<sub>2</sub>, 150ug ml<sup>-1</sup> BSA and 25ng of human genomic DNA of  
15 unknown source. Each pair of oligonucleotides was tested against 12 random samples of genomic DNA and the products electrophoresed on a 4% agarose (Gibco-BRL) gel. The expected product of 95 base pairs was seen for both ASOs in 8 of the 12 DNAs indicating that these individuals are heterozygous for this polymorphism. Two of the DNAs amplified with only the 1AFC oligonucleotide and are thus homozygous for the allele with the  
20 cytidine at this position. Two of the DNAs amplified with only the 1AFT oligonucleotide and are thus homozygous for the allele with the thymine at this position.

PSP1 oligonucleotides 1BFC, 1BFT and 1BR were designed for detecting the polymorphism at nucleotide 1435 (cytidine to thymine) causing the Ala to Val amino acid change.

25 1BFC: TGG CGG GCT TTG GGG GGC ATT C 22mer (SEQ ID NO:35)

1BFT: TGG CGG GCT TTG GGG GGC ATT T 22mer (SEQ ID NO:36)

1BR: GAC GTC AGC AGG GCC CGG AGG TC 23mer (SEQ ID NO:37)



発明番号 156220

Pairs of oligonucleotides (1BFC + 1BR, or 1BFT + 1BR) were used in a PCR under the following conditions: 94°C for 40 seconds, 67°C for 30 seconds, for 35 cycles in a reaction containing 1 U KlenTaqI (GenPak Ltd.), 50mM Tris-Cl pH9.1, 16mM ammonium sulphate, 3.5mM MgCl<sub>2</sub>, 150ug ml<sup>-1</sup> BSA and 25ng of human genomic DNA of unknown source. Each pair of oligonucleotides was tested against 12 random samples of genomic DNA and the products electrophoresed on a 4% agarose (Gibco-BRL) gel. The expected product of 75 base pairs was seen using the 1BFT ASO in 9 of the 12 samples indicating that the other 3 individuals have a different allele at this position.

#### 10 Example 5 - Detecting the D87258 polymorphism

Oligonucleotides 2AFG, 2AFT and 2AR were designed for detecting the polymorphism at nucleotide 1325 (guanine to thymine) causing the Gly to Val amino acid change.

2AFG: GAT ACC CCA GCA GAA GCT GG 20mer (SEQ ID NO:38)

15 2AFT: GAT ACC CCA GCA GAA GCT GT 20mer (SEQ ID NO:39)

2AR: GCT GAC ATC ATT GGC GGA GAC 21mer (SEQ ID NO:40)

Pairs of oligonucleotides (2AFG + 2AR, or 2AFT + 2AR) were used in a PCR under the following conditions: 94°C for 40 seconds, 62°C for 30 seconds, for 35 cycles in a reaction containing 1 U KlenTaqI (GenPak Ltd.), 50mM Tris-Cl pH9.1, 16mM ammonium sulphate, 3.5mM MgCl<sub>2</sub>, 150ug ml<sup>-1</sup> BSA and 25ng of human genomic DNA of unknown source. Each pair of oligonucleotides was tested against 12 random samples of genomic DNA and the products electrophoresed on a 4% agarose (Gibco-BRL) gel. The 2AFT ASO generated a band of approximately 1000 bp. The predicted band was 90 bp.

25 Presumably, the presence of the larger bands was due to the presence of an intron in the region flanked by oligonucleotides 2AR and 2AFT. Bands were observed in all of the samples amplified with 2AFT indicating that the allele containing the thymine is present in all 12 individuals.

30 The present invention may be embodied in other specific forms without departing from the spirit or essential attributes thereof, and, accordingly, reference should be made to the appended claims, rather than to the foregoing specification, as indicating the scope of the invention.

整理番号 156220

## SEQ ID NO:1:

```

GGGACTCCCC CAAACCAATG TGAATACAT TCAAACTGAT GCAGCTATTG ATTTTGAAA      60
CTCTGGAGGT CCCCTGGTTA ACCTGGATGG GGAGGTGATT GGAGTGAACA CCATGAAGGT      120
CACAGCTGGA ATCTCCTTTG CCATCCCTTC TGATCGTCTT CGAGAGTTTC TGCATCGTGG      180
GGAAAAAAG AATTCTCTCT CCGCAATCAG TGGGTCCCAG CGGCGCTACA TTGGGGTGAT      240
GATGCTGACC CTGAGTCCCA GCATCCTTGC TGAACACAG CTTCGAGAAC CAAGCTTTCC      300
CGATGTTTCA CATGGTGTAC TCATCCATAA AGTCATCCTG GGCTCCCCTG CACACCGGGC      360
TGGTCTGCGG CCTGGTGATG TGATTTTGGC CATTGGGGAG CAGATGGTAC AAAATGCTGA      420
AGATGTTTTAT GAAGCTGTTT GAACCCAATC CCAGTTGGCA GTGCAGATCC GCGGGGGACG      480
AGAAACACTG ACCTTATATG TGACCCCTGA GGTACACAGAA TGAATAGATC ACCAAGACTA      540
TGAGGCTCCT GCTCTGATTT CCTCCTTGCC TTTCTGGCTG AGGTTCTGAG GGCACCGAGA      600
CAGAGGGTTA AATGAACCAG TGGGGCCAGG TCCCTCCAAC CACCAGCACT GACTCCTGGG      660
CTCTGAAGAA TCACAGAAAC ACTTTTATA TAAATAAAA TTATACCTAG CAACAAAAAA      720
AAAAAAAAAA AA                                     732

```

## SEQ ID NO:2:

```

Gly Leu Pro Gln Thr Asn Val Glu Tyr Ile Gln Thr Asp Ala Ala Ile
 1           5           10          15
Asp Phe Gly Asn Ser Gly Gly Pro Leu Val Asn Leu Asp Gly Glu Val
 20          25          30
Ile Gly Val Asn Thr Met Lys Val Thr Ala Gly Ile Ser Phe Ala Ile
 35          40          45
Pro Ser Asp Arg Leu Arg Glu Phe Leu His Arg Gly Glu Lys Lys Asn
 50          55          60
Ser Ser Ser Gly Ile Ser Gly Ser Gln Arg Arg Tyr Ile Gly Val Met
 65          70          75          80
Met Leu Thr Leu Ser Pro Ser Ile Leu Ala Glu Leu Gln Leu Arg Glu
 85          90          95
Pro Ser Phe Pro Asp Val Gln His Gly Val Leu Ile His Lys Val Ile
100         105         110
Leu Gly Ser Pro Ala His Arg Ala Gly Leu Arg Pro Gly Asp Val Ile
115         120         125
Leu Ala Ile Gly Glu Gln Met Val Gln Asn Ala Glu Asp Val Tyr Glu
130         135         140
Ala Val Arg Thr Gln Ser Gln Leu Ala Val Gln Ile Arg Arg Gly Arg
145         150         155         160
Glu Thr Leu Thr Leu Tyr Val Thr Pro Glu Val Thr Glu
165         170

```

## SEQ ID NO:3:

```

CCCAGTCTCT GGGCCCGGTT GTCTGTTGGG GTCAGTGAAC CCCGAGCATG CCTGACGTCT      60
GGGACCCCGG GTCCCCGGGC ACAACTGACT GCGGTGACCC CAGATACCAG GACCCGGGAG      120

```

整理番号 156220

GCCTCAGACA ACTCTGGAAC COGTTGCGGC GCGTGGCTGG CCGTGGCGCT GGGCGCTGGG 180  
 GGGGCAGTGC TGTGTGTGTT GTGGGGCGGG GGTGCGGGTC CTCGGGCCGT CCTCGCCGCC 240  
 GTCCCTAGCC CGCCGCCCGC TTCTCCCCGG AGTCAGTACA ACTTCATCGC AGATGTGGTG 300  
 GAGAAGACAG CACCTGOCGT GGTCTATATC GAGATCCTGG ACCGGCACOC TTCTTTGGGC 360  
 CGCGAGGTCC CTATCTCGAA CGGCTCAGGA TTCGTGGTGG CTGCOGATGG GCTCATTGTC 420  
 ACCAAGGCC ATGTGGTGGC TGATCGGCGC AGAGTCCGTG TGAGACTGCT AAGCGGCGAC 480  
 ACGTATGAGG CCGTGGTCAC AGCTGTGGAT CCCGTGGCAG ACATCGCAAC GCTGAGGATT 540  
 CAGACTAAGG AGCCTCTCCC CACGCTGCCT CTGGGACGCT CAGCTGATGT CCGGCAAGGG 600  
 GAGTTTGTG TTGCCATGGG AAGTCCCTTT GCACTGCAGA ACACGATCAC ATCCGGCATT 660  
 GTTAGCTCTG CTCAGCGTCC AGCCAGAGAC CTGGGACTCC CCCAAACCAA TGTGGAATAC 720  
 ATTCAAATG ATGCAGCTAT TGATTTTGA AACTCTGGAG GTCCCTGGT TAACCTGGTG 780  
 AGTGAGACAT CCTTCCTOC AAGAATCCCT GCCCCAGGTC AGTGTGGGAA GGGTAGGTTT 840  
 CCCCIAATC AAGGATGTTT GGTCAGTTT CTGAGCAGTT CTTTGTGGC TATCTCTCAA 900  
 TATCCAAACA GATCTCCCA AACTTTGCTG GTACTTTGT TCGGGTGCC CCATCCCTTA 960  
 CTATTGTITT AGGCTAGGGA ACTGGGGGCT GTATCCCTGC AGGATGGGGA GGTGATTGGA 1020  
 GTGAACACCA TGAAGGTGAC AGCTGGAATC TCCTTTGCCA TCCCTTCTGA TCGTCTTCGA 1080  
 GAGTTTCTGC ATCGTGGGGA AAAGAAGRAAT TCCTCTCTCG GAATCAGTGG GTCCAGCGG 1140  
 CGCTACATTG GGGTGATGAT GCTGACCCTG AGTCCAGCA TCCTTGCTGA ACTACAGCTT 1200  
 CGAGAACCAA GCTTTCOCGA TGTTGAGCAT GGTGTACTCA TCCATAAAGT CATCCTGGGC 1260  
 TCCCCGACAC ACCGGGCTGG TCTGCGGCCT GGTGATGTGA TTTTGGCCAT TGGGGAGCAG 1320  
 ATGGTACAAA ATGCTGAAGA TGTTATGAA GCTGTTGAA CCCAATCCCA GTTGGCAGTG 1380  
 CAGATCGGC GGGGACGAGA AACACTGACC TTATATGTGA CCGCTGAGGT CACAGAATGA 1440  
 ATAGATCACC AAGAGTATGA GGCTCCTGCT CTGATTCTCT CCTTGCCCTT CTGGCTGAGG 1500  
 TTCTGAGGGC ACCGAGACAG AGGGTTAAAT GAACAGTGG GGGCAGGTCC CTCCAACCAC 1560  
 CAGCACTGAC TCCTGGGCTC TGAAGAATCA CAGAAACACT TTTTATATAA AATRAAATTA 1620  
 TACCTAGCAA CATATTATAG TAAAAAATGA GGTGGGAGGG CTGGATCTTT TCCCCACCA 1680  
 AARGGCTAGA GGTAAGCTG TATCCCCCTA AACTTAGGGG AGATACTGGA GCTGACCATC 1740  
 CTGACCTCCT ATTAAGAAA ATGAGCTGCT GAAAAA AAAA 1787

SBQ ID NO:4:

Pro Ser Leu Trp Ala Arg Leu Ser Val Gly Val Thr Glu Pro Arg Ala  
 1 5 10 15  
 Cys Leu Thr Ser Gly Thr Pro Gly Pro Arg Ala Gln Leu Thr Ala Val  
 20 25 30  
 Thr Pro Asp Thr Arg Thr Arg Glu Ala Ser Glu Asn Ser Gly Thr Arg  
 35 40 45  
 Ser Arg Ala Trp Leu Ala Val Ala Leu Gly Ala Gly Gly Ala Val Leu  
 50 55 60  
 Leu Leu Leu Trp Gly Gly Gly Arg Gly Pro Pro Ala Val Leu Ala Ala  
 65 70 75 80  
 Val Pro Ser Pro Pro Pro Ala Ser Pro Arg Ser Gln Tyr Asn Phe Ile  
 85 90 95

整理番号 156220

Ala Asp Val Val Glu Lys Thr Ala Pro Ala Val Val Tyr Ile Glu Ile  
 100 105 110  
 Leu Asp Arg His Pro Phe Leu Gly Arg Glu Val Pro Ile Ser Asn Gly  
 115 120 125  
 Ser Gly Phe Val Val Ala Ala Asp Gly Leu Ile Val Thr Asn Ala His  
 130 135 140  
 Val Val Ala Asp Arg Arg Arg Val Arg Val Arg Leu Leu Ser Gly Asp  
 145 150 155 160  
 Thr Tyr Glu Ala Val Val Thr Ala Val Asp Pro Val Ala Asp Ile Ala  
 165 170 175  
 Thr Leu Arg Ile Gln Thr Lys Glu Pro Leu Pro Thr Leu Pro Leu Gly  
 180 185 190  
 Arg Ser Ala Asp Val Arg Gln Gly Glu Phe Val Val Ala Met Gly Ser  
 195 200 205  
 Pro Phe Ala Leu Gln Asn Thr Ile Thr Ser Gly Ile Val Ser Ser Ala  
 210 215 220  
 Gln Arg Pro Ala Arg Asp Leu Gly Leu Pro Gln Thr Asn Val Glu Tyr  
 225 230 235 240  
 Ile Gln Thr Asp Ala Ala Ile Asp Phe Gly Asn Ser Gly Gly Pro Leu  
 245 250 255  
 Val Asn Leu Val Ser Glu Thr Ser Phe Leu Pro Arg Ile Pro Ala Pro  
 260 265 270  
 Gly Gln Cys Gly Lys Gly Arg Phe Pro Leu Ile Gln Gly Cys Leu Val  
 275 280 285  
 Lys Phe Leu Ser Ser Ser Leu Ala Ile Ser Gln Tyr Pro Thr Arg  
 290 295 300  
 Ser Pro Gln His Leu Leu Val Leu Leu Phe Gly Cys Pro His Pro Leu  
 305 310 315 320  
 Leu Phe Val

SEQ ID NO:5:

CTTCGGGCA7 GCGGGGCTTT GGGGGGCATT CGCTGGGGGA GGAGACCCCG TTTGACCCCT 60  
 GACCTCCGGG CCCTGCTGAC GTCAGGAAT TCTGACCCCG GGGCCCGAGT GACTTATGGG 120  
 ACCCCAGATC TCTGGGCCCC GTTGTCTGTT GGGGTCCTG AACCCCGAGC ATGCCTGACG 180  
 TCTGGGACCC CGGGTCCCCG GGCACAAC TG ACTGCGGTGA CCCAGATAC CAGGACCCGG 240  
 GAGGCCTCAG AGAATCTTGG AACCCGTTCC CGCGCGTGCC TGGCGGTGGC GCTGGGCGCT 300  
 GGGGGGGCAG TGCTGTGTGT GTTGTGGGGC GGGGGTCGGG GTCCTCCGGC CGTCCTCGCC 360  
 GCCGTCCCTA GCGCGCCGCC CGCTTCTCCC CGGAGTCAGT ACAACTTCAT CGCAGATGTG 420  
 GTGGAGAAGA CAGCACCTGC CGTGGTCTAT ATCGAGATCC TGGACCGGCA CCCTTTCTTG 480  
 GGCCGCGAGG TCCCTATCTC GAACGGCTCA GGATTCGTGG TGGCTGCCGA TGGGCTCATT 540  
 GTCACCAACG CCCATGTGGT GGCTGATCGG CGCAGAGTCC GTGTGAGACT GCTAAGCGGC 600  
 GACACGTATG AGGCCGTGGT CACACCTGTG GATCCCGTGG CAGACATCGC AACGCTGAGG 660

整理番号 156220

```

ATTGAGACTA AGGAGCCTCT CCCACGCTG CCTCTGGGAC GTCAGCTGA TCTCCGGCAA 720
GGGGAGTTTG TTGTTGCCAT GGGAGTCCC TTIGCACTGC AGAACACGAT CACATCCGGC 780
ATTGTTAGCT CTGCTCAGCG TCCAGCCAGA GACCTGGGAC TCCCCAAAC CAATGTGGAA 840
TACATTCAAA CTGATGCAGC TATTGATTTT GGAACTCTG GAGGTCCCCT GGTTAACCTG 900
GCTAGGGAAC TCGGGCCTGT ATCCCTGCAG GATGGGGAGG TGATTGGAGT GAACACCATG 960
AAGGTCACAG CTGGAACTC CTTTGCCATC CCTTCTGATC GTCTTCGAGA GTTCTCCAT 1020
CGTGGGGAAA AGAAGAATTC CTCCTCCGGA ATCAGTGGGT CCCAGCGGCG CTACATTGGG 1080
GTGATGATGC TGACCTGAG TCCAGGGCT GGTCTGCGGC CTGGTGATGT GATTTTGGCC 1140
ATTGGGGAGC AGATGGTACA AAATGCTGAA GATGTTTATG AAGCTGTTCT AACCCATCC 1200
CAGTGGCAG TGCAGATCCG GCGGGGACGA GAAACTCTGA CCTTATATCT GACCCCTGAG 1260
GTCACAGAAT CAATAGATCA CCAAGAGTAT GAGGCTCTG CTCTGATTTC CTCCTTGCTT 1320
TTCTGGCTGA GGTCTGAGG GCACCGAGAC AGAGGGTTAA ATGAACCAGT GGGGGCAGGT 1380
CCCTCCAAAC ACCAGCACTG ACTCCTGGGC TCTGAAGAAT CACAGAAACA CTTTTATAT 1440
AAAAIAAAAT TATCCTAGC AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 1500
AAA 1503

```

SEQ ID NO:6:

```

Leu Arg Ala Trp Arg Ala Leu Gly Gly Ile Arg Trp Gly Arg Arg Pro
1      5      10      15
Arg Leu Thr Pro Asp Leu Arg Ala Leu Leu Thr Ser Gly Thr Ser Asp
20      25      30
Pro Arg Ala Arg Val Thr Tyr Gly Thr Pro Ser Leu Trp Ala Arg Leu
35      40      45
Ser Val Gly Val Thr Glu Pro Arg Ala Cys Leu Thr Ser Gly Thr Pro
50      55      60
Gly Pro Arg Ala Gln Leu Thr Ala Val Thr Pro Asp Thr Arg Thr Arg
65      70      75      80
Glu Ala Ser Glu Asn Ser Gly Thr Arg Ser Arg Ala Trp Leu Ala Val
85      90      95
Ala Leu Gly Ala Gly Gly Ala Val Leu Leu Leu Trp Gly Gly Gly
100     105     110
Arg Gly Pro Pro Ala Val Leu Ala Ala Val Pro Ser Pro Pro Pro Ala
115     120     125
Ser Pro Arg Ser Gln Tyr Asn Phe Ile Ala Asp Val Val Glu Lys Thr
130     135     140
Ala Pro Ala Val Val Tyr Ile Glu Ile Leu Asp Arg His Pro Phe Leu
145     150     155     160
Gly Arg Glu Val Pro Ile Ser Asn Gly Ser Gly Phe Val Val Ala Ala
165     170     175
Asp Gly Leu Ile Val Thr Asn Ala His Val Val Ala Asp Arg Arg Arg
180     185     190
Val Arg Val Arg Leu Leu Ser Gly Asp Thr Tyr Glu Ala Val Val Thr

```

整理番号 156220

195	200	205
Ala Val Asp Pro Val Ala Asp Ile Ala Thr Leu Arg Ile Gln Thr Lys		
210	215	220
Glu Pro Leu Pro Thr Leu Pro Leu Gly Arg Ser Ala Asp Val Arg Gln		
225	230	235
Gly Glu Phe Val Val Ala Met Gly Ser Pro Phe Ala Leu Gln Asn Thr		
245	250	255
Ile Thr Ser Gly Ile Val Ser Ser Ala Gln Arg Pro Ala Arg Asp Leu		
260	265	270
Gly Leu Pro Gln Thr Asn Val Glu Tyr Ile Gln Thr Asp Ala Ala Ile		
275	280	285
Asp Phe Gly Asn Ser Gly Gly Pro Leu Val Asn Leu Ala Arg Glu Leu		
290	295	300
Gly Ala Val Ser Leu Gln Asp Gly Glu Val Ile Gly Val Asn Thr Met		
305	310	315
Lys Val Thr Ala Gly Ile Ser Phe Ala Ile Pro Ser Asp Arg Leu Arg		
325	330	335
Glu Phe Leu His Arg Gly Glu Lys Lys Asn Ser Ser Ser Gly Ile Ser		
340	345	350
Gly Ser Gln Arg Arg Tyr Ile Gly Val Met Met Leu Thr Leu Ser Pro		
355	360	365
Arg Ala Gly Leu Arg Pro Gly Asp Val Ile Leu Ala Ile Gly Glu Gln		
370	375	380
Met Val Gln Asn Ala Glu Asp Val Tyr Glu Ala Val Arg Thr Gln Ser		
385	390	395
Gln Leu Ala Val Gln Ile Arg Arg Gly Arg Glu Thr Leu Thr Leu Tyr		
405	410	415
Val Thr Pro Glu Val Thr Glu		
420		

SEQ ID NO:7:

GGCCGGAAGG GCTAGCGGTC CCAGCATACC CGCGGGCCCC TTGGGCGGTC TCACAACTCG	60
CGTCCGGCGG AGACCACAA7 TCCCGGCATT CGTGGGGCAT GGAGGAGTCG GCCTCCCGGA	120
ATCCTGGTCC CGGCGTGCCAC TTCTGAAGGA CTTCAAGTAC CGGCGTGCCC CGCGTCCTAC	180
TGTCCGCGTG CTCGCGTCC7 GGGTGCCGCC TCTGAGTAGG GCGGGCGAGG AGGCAGCCAA	240
GGCGGAGCTG ATG GCT GCG CCG AGG GCG GGG CGG GGT GCA GGC TGG AGC	289
Met Ala Ala Pro Arg Ala Gly Arg Gly Ala Gly Trp Ser	
1 5 10	
CTT CGG GCA TGG CGG GCT TTG GGG GGC ATT TGC TGG GGG AGG AGA CCC	337
Leu Arg Ala Trp Arg Ala Leu Gly Gly Ile Cys Trp Gly Arg Arg Pro	
15 20 25	

整理番号 156220

CGT TTG ACC CCT GAC CTC CGG GCC CTG CTG ACG TCA GGA ACT TCT GAC Arg Leu Thr Pro Asp Leu Arg Ala Leu Leu Thr Ser Gly Thr Ser Asp 30 35 40 45	385
CCC CGG GCC CGA GTG ACT TAT GGG ACC CCC AGT CTC TGG GCC CGG TTG Pro Arg Ala Arg Val Thr Tyr Gly Thr Pro Ser Leu Trp Ala Arg Leu 30 35 60	433
TCT GTT GGG GTC ACT GAA CCC CGA GCA TGC CTG ACG TCT GGG ACC CCG Ser Val Gly Val Thr Glu Pro Arg Ala Cys Leu Thr Ser Gly Thr Pro 65 70 75	481
GGT CCC CGG GCA CAA CTG ACT GCG CTG ACC CCA CAT ACC AGG ACC CGG Gly Pro Arg Ala Gln Leu Thr Ala Val Thr Pro Asp Thr Arg Thr Arg 80 85 90	529
GAG GCC TCA GAG AAC TCT GGA ACC CGT TCG CGC GCG TGG CTG GCG GTG Glu Ala Ser Glu Asn Ser Gly Thr Arg Ser Arg Ala Trp Leu Ala Val 95 100 105	577
GCG CTG GCG GCT CCG CGG GCA GTG CTG TTC TTC TTC TGC CCG CCG GCT Ala Leu Gly Ala Gly Gly Ala Val Leu Leu Leu Leu Trp Gly Gly Gly 110 115 120 125	625
CGG GGT CCT CCG GCC GTC CTC GCG GCC GTC CCT AGC CCG CCG CCC GCT Arg Gly Pro Pro Ala Val Leu Ala Ala Val Pro Ser Pro Pro Pro Ala 130 135 140	673
TCT CCC CGG AGT CAG TAC AAC TTC ATC GCA GAT GTG GTG GAG AAG ACA Ser Pro Arg Ser Gln Tyr Asn Phe Ile Ala Asp Val Val Glu Lys Thr 145 150 155	721
GCA CCT GCC GTG GTC TAT ATC GAG ATC CTG GAC CGG CAC CCT TTC TTG Ala Pro Ala Val Val Tyr Ile Glu Ile Leu Asp Arg His Pro Phe Leu 160 165 170	769
GGC CGC GAG GTC CCT ATC TCG AAC GGC TCA GGA TTC GTG GTG GCT GCC Gly Arg Glu Val Pro Ile Ser Asn Gly Ser Gly Phe Val Val Ala Ala 175 180 185	817
GAT GGG CTC ATT GTC ACC AAC GCC CAT GTG GTG GCT GAT CGG CGC AGA Asp Gly Leu Ile Val Thr Asn Ala His Val Val Ala Asp Arg Arg Arg 185 190 195	865

整理番号 156220

190	195	200	205	
GTC CGT GTG AGA CTG CTA AGC GGC GAC ACG TAT GAG GCC GTG GTC ACA Val Arg Val Arg Leu Leu Ser Gly Asp Thr Tyr Glu Ala Val Val Thr 210 215 220				913
GCT GTG GAT CCC GTG GCA GAC ATC GCA ACG CTG AGG ATT CAG ACT AAG Ala Val Asp Pro Val Ala Asp Ile Ala Thr Leu Arg Ile Gln Thr Lys 225 230 235				961
GAG CCT CTC CCC ACG CTG CCT CTG GGA CGC TCA GCT GAT GTC CGG CAA Glu Pro Leu Pro Thr Leu Pro Leu Gly Arg Ser Ala Asp Val Arg Gln 240 245 250				1009
GGG GAG TTT GTT GTT GCC ATG GGA AGT CCC TTT GCA CTG CAG AAC ACG Gly Glu Phe Val Val Ala Met Gly Ser Pro Phe Ala Leu Gln Asn Thr 255 260 265				1057
ATC ACA TCC GGC ATT GTT AGC TCT GCT CAG CGT CCA GCC AGA GAC CTG Ile Thr Ser Gly Ile Val Ser Ser Ala Gln Arg Pro Ala Arg Asp Leu 270 275 280 285				1105
GGA CTC CCC CAA ACC AAT GTC GAA TAC ATT CAA ACT GAT GCA GCT ATT Gly Leu Pro Gln Thr Asn Val Glu Tyr Ile Gln Thr Asp Ala Ala Ile 290 295 300				1153
GAT TTT GGA AAC TCT GGA GGT CCC CTG GTT AAC CTG GAT GGG GAG GTG Asp Phe Gly Asn Ser Gly Gly Pro Leu Val Asn Leu Asp Gly Glu Val 305 310 315				1201
ATT GGA GTG AAC ACC ATG AAG GTC ACA GCT GGA ATC TCC TTT GCC ATC Ile Gly Val Asn Thr Met Lys Val Thr Ala Gly Ile Ser Phe Ala Ile 320 325 330				1249
CCT TCT GAT CGT CTT CGA GAG TTT CTG CAT CGT GGG GAA AAG AAG AAT Pro Ser Asp Arg Leu Arg Glu Phe Leu His Arg Gly Glu Lys Lys Asn 335 340 345				1297
TCC TCC TCC GGA ATC AGT GGG TCC CAG CGG CGC TAC ATT GGG GTG ATG Ser Ser Ser Gly Ile Ser Gly Ser Gln Arg Arg Tyr Ile Gly Val Met 350 355 360 365				1345



登録番号 156220

ATG CTG ACC CTG AGT CCC AGC ATC CTT GCT GAA CTA CAG CTT CGA GAA Met Leu Thr Leu Ser Pro Ser Ile Leu Ala Glu Leu Gln Leu Arg Glu 370 375 380	1393
CCA AGC TTT CCC GAT GTT CAG CAT GGT GTA CTC ATC CAT AAA GTC ATC Pro Ser Phe Pro Asp Val Gln His Gly Val Leu Ile His Lys Val Ile 385 390 395	1441
CTG GGC TCC CCT GCA CAC CGG GCT GGT CTG CGG CCT GGT GAT GTG ATT Leu Gly Ser Pro Ala His Arg Ala Gly Leu Arg Pro Gly Asp Val Ile 400 405 410	1489
TTG GCC ATT GGG GAG CAG ATG GTA CAA AAT GCT GAA GAT GTT TAT GAA Leu Ala Ile Gly Glu Gln Met Val Gln Asn Ala Glu Asp Val Tyr Glu 415 420 425	1537
GCT GTT CGA ACC CAA TCC CAG TTG GCA GTG CAG ATC CGG CGG GGA CGA Ala Val Arg Thr Gln Ser Gln Leu Ala Val Gln Ile Arg Arg Gly Arg 430 435 440 445	1585
GAA ACA CTG ACC TTA TAT GTG ACC CCT GAG GTC ACA GAA TGAATAGATC ACC Glu Thr Leu Thr Leu Tyr Val Thr Pro Glu Val Thr Glu 450 455	1637
AAGAGTATGA GGCCTCTGCT CTGATTCCT CCTTGCCCTT CTGGCTGAGG TTCTGAGGGC ACCGAGACAG AGGGTTAAAT GAACCACTGG GGGCAGGTCC CTCCAACCAC CAGCACTGAC TCCTGGGCTC TGAAGAATCA CAGAAACACT TTTTATATAA AATAAAATTA TACCTAGCAA CATAAAAAAA AAAAAAAA	1697 1757 1817 1835

SEQ ID NO:8:

Met Ala Ala Pro Arg Ala Gly Arg Gly Ala Gly Trp Ser Leu Arg Ala 1 5 10 15
Trp Arg Ala Leu Gly Gly Ile Cys Trp Gly Arg Arg Pro Arg Leu Thr 20 25 30
Pro Asp Leu Arg Ala Leu Leu Thr Ser Gly Thr Ser Asp Pro Arg Ala 35 40 45
Arg Val Thr Tyr Gly Thr Pro Ser Leu Trp Ala Arg Leu Ser Val Gly 50 55 60
Val Thr Glu Pro Arg Ala Cys Leu Thr Ser Gly Thr Pro Gly Pro Arg 65 70 75 80
Ala Gln Leu Thr Ala Val Thr Pro Asp Thr Arg Thr Arg Glu Ala Ser 85 90 95

整理番号 156220

Glu Asn Ser Gly Thr Arg Ser Arg Ala Trp Leu Ala Val Ala Leu Gly  
 100 105 110  
 Ala Gly Gly Ala Val Leu Leu Leu Trp Gly Gly Gly Arg Gly Pro  
 115 120 125  
 Pro Ala Val Leu Ala Ala Val Pro Ser Pro Pro Pro Ala Ser Pro Arg  
 130 135 140  
 Ser Gln Tyr Asn Phe Ile Ala Asp Val Val Glu Lys Thr Ala Pro Ala  
 145 150 155 160  
 Val Val Tyr Ile Glu Ile Leu Asp Arg His Pro Phe Leu Gly Arg Glu  
 165 170 175  
 Val Pro Ile Ser Asn Gly Ser Gly Phe Val Val Ala Ala Asp Gly Leu  
 180 185 190  
 Ile Val Thr Asn Ala His Val Val Ala Asp Arg Arg Arg Val Arg Val  
 195 200 205  
 Arg Leu Leu Ser Gly Asp Thr Tyr Glu Ala Val Val Thr Ala Val Asp  
 210 215 220  
 Pro Val Ala Asp Ile Ala Thr Leu Arg Ile Gln Thr Lys Glu Pro Leu  
 225 230 235 240  
 Pro Thr Leu Pro Leu Gly Arg Ser Ala Asp Val Arg Gln Gly Glu Phe  
 245 250 255  
 Val Val Ala Met Gly Ser Pro Phe Ala Leu Gln Asn Thr Ile Thr Ser  
 260 265 270  
 Gly Ile Val Ser Ser Ala Gln Arg Pro Ala Arg Asp Leu Gly Leu Pro  
 275 280 285  
 Gln Thr Asn Val Glu Tyr Ile Gln Thr Asp Ala Ala Ile Asp Phe Gly  
 290 295 300  
 Asn Ser Gly Gly Pro Leu Val Asn Leu Asp Gly Glu Val Ile Gly Val  
 305 310 315 320  
 Asn Thr Met Lys Val Thr Ala Gly Ile Ser Phe Ala Ile Pro Ser Asp  
 325 330 335  
 Arg Leu Arg Glu Phe Leu His Arg Gly Glu Lys Lys Asn Ser Ser Ser  
 340 345 350  
 Gly Ile Ser Gly Ser Gln Arg Arg Tyr Ile Gly Val Met Met Leu Thr  
 355 360 365  
 Leu Ser Pro Ser Ile Leu Ala Glu Leu Gln Leu Arg Glu Pro Ser Phe  
 370 375 380  
 Pro Asp Val Gln His Gly Val Leu Ile His Lys Val Ile Leu Gly Ser  
 385 390 395 400  
 Pro Ala His Arg Ala Gly Leu Arg Pro Gly Asp Val Ile Leu Ala Ile  
 405 410 415  
 Gly Glu Gln Met Val Gln Asn Ala Glu Asp Val Tyr Glu Ala Val Arg  
 420 425 430

送呈請号 156220

Thr Gln Ser Gln Leu Ala Val Gln Ile Arg Arg Gly Arg Glu Thr Leu  
 435 440 445  
 Thr Leu Tyr Val Thr Pro Glu Val Thr Glu  
 450 455

SEQ ID NO:9:

TGGGACAGGC AGCTCOGGGG TCCGCGGTTT CACATCGGAA ACAAACAGC GGCTGGTCTG 60  
 GAAGGAACCT GAGCTACGAG CCGCGGCGGC AGCGGGGCGG CGGGGAAGCG TATACCTAAT 120  
 CTGGGAGCCT CCAAGTGACA ACAGCCTTTG CGGTCCCTTAG ACAGCTTGGC CTGGAGGAGA 180  
 ACACATGAAA GAAAGAACCT CAAGAGGCTT TGTTTTCTGT GAAACAGTAT TTCTATACAG 240  
 TTGCTCCAT GACAGAGTTA CCTGCACCGT TGTCCTACTT CCAGAATGCA CAGATGTCTG 300  
 AGGACAACCA CCTGAGCAAT ACTGTACGTA GCCAGAATGA CAATAGAGAA CGGCAGGAGC 360  
 ACAACGACAG ACGGAGCCTT GGCCACCCTG AGCCATTATC TAATGGACGA CCCCAGGGTA 420  
 ACTCCCGGCA GGTGGTGGAG CAAGATGAGG AAGAAGATGA GGAGCTGACA TTGAATATG 480  
 GCGCCAGCA TGTGATCATG CTCCTTGTC CTGTGACTCT CTGCATGGTG GTGGTCGTGG 540  
 CTACCATTA GTCAGTCAGC TTTTATACCC GGAAGGATGG GCAGCTAATC TATACCCCAT 600  
 TCACAGAAGA TACCGAGACT GTGGGCCAGA GAGCCCTGCA CTCATTTCTG AATGCTGCCA 660  
 TCATGATCAG TGTCATTGTT GTCATGACTA TCCTCCTGGT GGTCTGTAT AAATACAGGT 720  
 GCTATAAGGT CATCCATGCC TGGCTTATTA TATCATCTCT ATTGTGTCTG TTCTTTTTTT 780  
 CATTCAATTA CTTGGGGGAA GTGTTTAAAA CCTATAACGT TGCTGTGGAC TACATTACTG 840  
 TTGCACTCCT GATCTGGAAT TTTGGTGTGG TGGGAATGAT TTCCATTAC TGGAAAGGTC 900  
 CACTTCGACT CCAGCAGGCA TATCTCAATTA TGATTAGTGC CCTCATGGCC CTGGTGTTTA 960  
 TCAAGTACCT CCCTGAATGG ACTGCGTGGC TCATCTTGGC TGTGATTTCA GTATATGATT 1020  
 TAGTGGCTGT TTTGTGTCCG AAGGTCCAC TTCGTATGCT GGTGAAACA GCTCAGGAGA 1080  
 GAAATGAAAC GCTTTTCCA GCTCTCAAT ACTCCTCAAC AATGGTGTGG TTGGTGAATA 1140  
 TGGCAGAAGG AGACCCGGAA GCTCAAAGGA GAGTATCCAA AATTCCAAG TATATGCAG 1200  
 AAAGCACAGA AAGGGAGTCA CAAGACACTG TTGCAGAGAA TGATGATGGC GGGTTCAGTG 1260  
 AGGAATGGGA AGCCAGAGG GACAGTCATC TAGGGCCTCA TCGCTCTAQA CCTGAGTCAC 1320  
 GAGCTGCTGT CCAGGAACCT TCCAGCAGTA TCCTCGCTGG TGAAGACCCA GAGGAAAGGG 1380  
 GAGTAAACT TGGATTGGGA GATTTCAIT TCTACAGTGT TCTGGTTGGT AAAGCCTCAG 1440  
 CAACAGCCAG TGGAGACTGG AACACAACCA TAGCCTGTTT CGTAGCCATA TTAATTGGTT 1500  
 TGTGCCTTAC ATTATTACTC CTTGCCATTT TCAAGAAAGC ATTGCCAGCT CTTCATCT 1560  
 CCATCACCTT TGGGCTTGTT TTCTACTTTG CCACAGATTA TCTGTACAG CCTTTATGG 1620  
 ACCAATTAGC ATTCCATCAA TTTTATATCT AGCATATTG CGGTAGAAT CCCATGGATG 1680  
 TTTCTTCTTT GACTATAACC AAATCTGGG AGGACAAAGG TGATTTTCCT GTGTCCACAT 1740  
 CTAACAAAGT CAAGATTCCC GGCTGGACTT TTGCAGCTTC CTTCAGTGC TTCCTGACCA 1800  
 CCTTGCACTA TTGGACTTTG GAAGGAGGTG CCTATAGAAA ACGATTTTGA ACATACTTCA 1860  
 TCGCAGTGGG CTGTGTCCCT CGGTGCAGAA ACTACCAGAT TTGAGGGACG AGGTCAAGGA 1920  
 GATATGATAG GCCCGGAAGT TGCTGTGCCC CATCAGCAGC TTGACGCGTG GTCACAGGAC 1980  
 GATTTCACTG AACTGCGAA CTCACAGGAC TACCGGTTAC CAAGAGGTTA GGTGAAGTGG 2040  
 TTTAAACCAA ACGGAACCTCT TCATCTTAAA CTACACGTTG AAAATCAACC CAATAATTCT 2100  
 GTATTAACCT AATTCTGPAC TTTTCAGGAG GTACTGTGAG GAAGAGCAGG CACCAGCAGC 2160

整理番号 156220

AGAATGGGGA ATGGAGAGGT GGGCAGGGGT TCCAGCTTCC CTTTGATTTT TTGCTGCAGA 2220  
 CTCATCCTTT TTAATGAGA CTTGTTTTCC CCTCTCTTG AGTCAAGTCA AATATGTAGA 2280  
 TTGCCTTTGG CAATTCTTCT TCTCAAGCAC TGACACTCAT TACCGTCTGT GATTGCCATT 2340  
 TCTTCCCAAG GCCAGTCTGA ACCTGAGGTT GCTTTATCCT AAAAGTTTTA ACCTCAGGTT 2400  
 CCAAATTCAG TAAATTTTGG AAACAGTACA GCTATTCTC ATCAATTCTC TATCATGTTG 2460  
 AAGTCAAATT TGGATTTTCC ACCAAATTCT GAATTGTAG ACATACTTGT ACGCTCACTT 2520  
 GCGCCAGAT GCTCTCTCTG TCCTCATTCT TCTCTCCAC ACAAGCAGTC TTTTCTACA 2580  
 GCCAGTAAGG CAGCTCTGTC ATGGTAGCAG ATGGTCCAT TATTCTAGGG TCTTACTCTT 2640  
 TGTATGATGA AAGAATGTG TTATGAATCG GTGCTGTCAG CCCTGCTGTC AGACCTTCTT 2700  
 CCACAGCAAA TGAGATGTAT GCCCAAAGCG GTAGAATTAA AGAAGAGTAA AATGGCTGTT 2760  
 GAAG 2764

SEQ ID NO:10:

Met Thr Glu Leu Pro Ala Pro Leu Ser Tyr Phe Gln Asn Ala Gln Met  
 1 5 10 15  
 Ser Glu Asp Asn His Leu Ser Asn Thr Val Arg Ser Gln Asn Asp Asn  
 20 25 30  
 Arg Glu Arg Gln Glu His Asn Asp Arg Arg Ser Leu Gly His Pro Glu  
 35 40 45  
 Pro Leu Ser Asn Gly Arg Pro Gln Gly Asn Ser Arg Gln Val Val Glu  
 50 55 60  
 Gln Asp Glu Glu Glu Asp Glu Glu Leu Thr Leu Lys Tyr Gly Ala Lys  
 65 70 75 80  
 His Val Ile Met Leu Phe Val Pro Val Thr Leu Cys Met Val Val Val  
 85 90 95  
 Val Ala Thr Ile Lys Ser Val Ser Phe Tyr Thr Arg Lys Asp Gly Gln  
 100 105 110  
 Leu Ile Tyr Thr Pro Phe Thr Glu Asp Thr Glu Thr Val Gly Gln Arg  
 115 120 125  
 Ala Leu His Ser Ile Leu Asn Ala Ala Ile Met Ile Ser Val Ile Val  
 130 135 140  
 Val Met Thr Ile Leu Leu Val Val Leu Tyr Lys Tyr Arg Cys Tyr Lys  
 145 150 155 160  
 Val Ile His Ala Trp Leu Ile Ile Ser Ser Leu Leu Leu Phe Phe  
 165 170 175  
 Phe Ser Phe Ile Tyr Leu Gly Glu Val Phe Lys Thr Tyr Asn Val Ala  
 180 185 190  
 Val Asp Tyr Ile Thr Val Ala Leu Leu Ile Trp Asn Phe Gly Val Val  
 195 200 205  
 Gly Met Ile Ser Ile His Trp Lys Gly Pro Leu Arg Leu Gln Gln Ala  
 210 215 220

登録番号 156220

Tyr Leu Ile Met Ile Ser Ala Leu Met Ala Leu Val Phe Ile Lys Tyr  
 225 230 235 240  
 Leu Pro Glu Trp Thr Ala Trp Leu Ile Leu Ala Val Ile Ser Val Tyr  
 245 250 255  
 Asp Leu Val Ala Val Leu Cys Pro Lys Gly Pro Leu Arg Met Leu Val  
 260 265 270  
 Glu Thr Ala Gln Glu Arg Asn Glu Thr Leu Phe Pro Ala Leu Ile Tyr  
 275 280 285  
 Ser Ser Thr Met Val Trp Leu Val Asn Met Ala Glu Gly Asp Pro Glu  
 290 295 300  
 Ala Gln Arg Arg Val Ser Lys Asn Ser Lys Tyr Asn Ala Glu Ser Thr  
 305 310 315 320  
 Glu Arg Glu Ser Gln Asp Thr Val Ala Glu Asn Asp Asp Gly Gly Phe  
 325 330 335  
 Ser Glu Glu Trp Glu Ala Gln Arg Asp Ser His Leu Gly Pro His Arg  
 340 345 350  
 Ser Thr Pro Glu Ser Arg Ala Ala Val Gln Glu Leu Ser Ser Ser Ile  
 355 360 365  
 Leu Ala Gly Glu Asp Pro Glu Glu Arg Gly Val Lys Leu Gly Leu Gly  
 370 375 380  
 Asp Phe Ile Phe Tyr Ser Val Leu Val Gly Lys Ala Ser Ala Thr Ala  
 385 390 395 400  
 Ser Gly Asp Trp Asn Thr Thr Ile Ala Cys Phe Val Ala Ile Leu Ile  
 405 410 415  
 Gly Leu Cys Leu Thr Leu Leu Leu Leu Ala Ile Phe Lys Lys Ala Leu  
 420 425 430  
 Pro Ala Leu Pro Ile Ser Ile Thr Phe Gly Leu Val Phe Tyr Phe Ala  
 435 440 445  
 Thr Asp Tyr Leu Val Gln Pro Phe Met Asp Gln Leu Ala Phe His Gln  
 450 455 460  
 Phe Tyr Ile  
 465

SEQ ID NO:11:

CGGAATTCGG TATGCTGGTT GAAACA

26

SEQ ID NO:12:

CGGGATCCTC AGGCTACGAA ACAGGCTAT

29

SEQ ID NO:13:

TATATCAGCG GTATGACCGA CCTCTATGCG TGGGATGAAT ACCGACGTCT GATGGCCGTA  
 GAACAATAAC CAGGCTTTTG TAAAGACGAA CAATAAATTT TTACCTTTTG CAGAACTTT

60

120

整理番号 156220

```

AGTTCGGAAC TTCAGGCTAT AAAACGAATC TGAAGAACAC AGCAATTTTG CGTTATCTGT 180
TAATCGAGAC TGAAATACAT GAAAAAACC ACATTAGCAC TGAGTCGACT GGCTCTGAGT 240
TTAGGTTTGG CGTTATCTCC GCTCTCTGCA ACGGCGGCTG AGACTTCTTC AGCAACGACA 300
GCCCAGCAGA TGCCAAGCCT TGCACCGATG CTCGAAAAGG TGATGCCTTC AGTGGTCAGC 360
ATTAAOCTAG AAGGTAGCAC AACCGTTAAT ACGCCGCGTA TGCCGCGTAA TTTCCAGCAC 420
TTCTTOGGTG ATGATTCTCC GTTCTGOCAG GAAGGTTCTC CGTTCAGAG CTCTCCGTTT 480
TGCCAGGGTG GCCAGGGCGG TAATGGTGGC GGCCAGCAAC AGAAATTCAT GCGCTGGGT 540
TCCGGCGTCA TCATTGATGC CGATAAAGGC TATGTCGTCA CCAACAACCA CGTGTGTGAT 600
AACGCGACGG TCATTAAAGT TCAACTGAGC GATGGCCGTA AGTTCGACGC GAAGATGGTT 660
GGCAAAGATC CGCGCTCTGA TATCGCGCTG ATCCAAATCC AGAACCAGAA AAACCTGACC 720
GCAATTAAGA TGGCGGAATC TGATGCACTG CGCGTGGGTG ATTACACCGT AGGGATTGGT 780
AACCCGTTTG GTCTGGGCGA GACGGTAATC TCCGGGATTG TCTCTGOGCT GGGGCGTAGC 840
GGCCTGAATG CCGAAAACTA CGAAAACCTC ATCCAGACCG ATGCAGCGAT CAACCGTGGT 900
AACTCCGGTG GTGCGCTGGT TAACCTGAAC GGCGAACTGA TCGGTATCAA CACCGCGATC 960
CTGCAACGGG ACGGCGGCAA CATCGGTATC GGTTTTGCTA TCCCGAGTAA CATGGTGAAA 1020
AACCTGACCT CGCAGATGGT GGAATACGGC CAGGTGAAAC CGGTGAGCT GGGTATTATG 1080
GGGACTGAGC TGAATCCGA ACTGGCGAA GCGATGAAAG TTGACGCCCA GCGCGGTGCT 1140
TTGTAAGCC AGGTTCTGCC TAATTCCTCC GCTGCAAAAG CGGCAATTAA AGCGGGTGAT 1200
GTGATCACCT CACTGAACGG TAAGCCGATC AGCAGCTTTG CGGCACTGCG TGCTCAGGTG 1260
GGTACTATGC CGGTAGGCGG CAAACTGACC CTGGGCTTAC TGCAGCAGCG TAAGCAGGTT 1320
AACGTGAACC TGGAACTGCA GCAGAGCAGC CAGAATCAGG TTGATTCCAG CTCCATCTTC 1380
AACGGCATTG AAGGCGCTGA GATGAGCAAC AAAGGCAAAG ATCAGGGCGT GGTAGTGAAC 1440
AACGTGAAA CCGGCACTCC CGCTGCGCAG ATCGGCTGTA AGAAAGGTGA TGTGATTATT 1500
GGCGCGAACC AGCAGGCAGT GAAAACATC GCTGAATGCT GTAAAGTTCT CGACAGCAAA 1560
CGTCTGTGTC TGGCACTCAA CATTGAGCGC GCGGACCGCC ATCTACCTGT TAATGCASTA 1620
ATCTCCCTCA ACCCCTTCCT GAAAACGGGA AGGGGTTCTC CTTACAATCT GTGAACCTCA 1680
CCACAATCC ATACATCTTC ATCATCCTTT AGGCATTGTC ACAATGCCGT ACGTTACGTA 1740
CTTCTTATG CTAAGCCGTC CATAACGGAG GACTTATGGC TGGCTGGCAT CTGATACCA 1800
APATGGCGCA GGATATCGTG GCACGTACCA TGGCGATCAT CGATACCAAT ATCA 1854

```

SEQ ID NO:14:

```

Met Lys Lys Thr Thr Leu Ala Leu Ser Arg Leu Ala Leu Ser Leu Gly
 1           5           10          15
Leu Ala Leu Ser Pro Leu Ser Ala Thr Ala Ala Glu Thr Ser Ser Ala
 20          25          30
Thr Thr Ala Gln Gln Met Pro Ser Leu Ala Pro Met Leu Glu Lys Val
 35          40          45
Met Pro Ser Val Val Ser Ile Asn Val Glu Gly Ser Thr Thr Val Asn
 50          55          60
Thr Pro Arg Met Pro Arg Asn Phe Gln Gln Phe Phe Gly Asp Asp Ser
 65          70          75          80
Pro Phe Cys Gln Glu Gly Ser Pro Phe Gln Ser Ser Pro Phe Cys Gln

```

整理番号 156220

85	90	95
Gly Gly Gln Gly Gly Asn Gly Gly Gly Gln Gln Gln Lys Phe Met Ala		
100	105	110
Leu Gly Ser Gly Val Ile Ile Asp Ala Asp Lys Gly Tyr Val Val Thr		
115	120	125
Asn Asn His Val Val Asp Asn Ala Thr Val Ile Lys Val Gln Leu Ser		
130	135	140
Asp Gly Arg Lys Phe Asp Ala Lys Met Val Gly Lys Asp Pro Arg Ser		
145	150	155
Asp Ile Ala Leu Ile Gln Ile Gln Asn Pro Lys Asn Leu Thr Ala Ile		
165	170	175
Lys Met Ala Asp Ser Asp Ala Leu Arg Val Gly Asp Tyr Thr Val Gly		
180	185	190
Ile Gly Asn Pro Phe Gly Leu Gly Glu Thr Val Thr Ser Gly Ile Val		
195	200	205
Ser Ala Leu Gly Arg Ser Gly Leu Asn Ala Glu Asn Tyr Glu Asn Phe		
210	215	220
Ile Gln Thr Asp Ala Ala Ile Asn Arg Gly Asn Ser Gly Gly Ala Leu		
225	230	235
Val Asn Leu Asn Gly Glu Leu Ile Gly Ile Asn Thr Ala Ile Leu Ala		
245	250	255
Pro Asp Gly Gly Asn Ile Gly Ile Gly Phe Ala Ile Pro Ser Asn Met		
260	265	270
Val Lys Asn Leu Thr Ser Gln Met Val Glu Tyr Gly Gln Val Lys Arg		
275	280	285
Gly Glu Leu Gly Ile Met Gly Thr Glu Leu Asn Ser Glu Leu Ala Lys		
290	295	300
Ala Met Lys Val Asp Ala Gln Arg Gly Ala Phe Val Ser Gln Val Leu		
305	310	315
Pro Asn Ser Ser Ala Ala Lys Ala Gly Ile Lys Ala Gly Asp Val Ile		
325	330	335
Thr Ser Leu Asn Gly Lys Pro Ile Ser Ser Phe Ala Ala Leu Arg Ala		
340	345	350
Gln Val Gly Thr Met Pro Val Gly Ser Lys Leu Thr Leu Gly Leu Leu		
355	360	365
Arg Asp Gly Lys Gln Val Asn Val Asn Leu Glu Leu Gln Gln Ser Ser		
370	375	380
Gln Asn Gln Val Asp Ser Ser Ser Ile Phe Asn Gly Ile Glu Gly Ala		
385	390	395
Glu Met Ser Asn Lys Gly Lys Asp Gln Gly Val Val Val Asn Asn Val		
405	410	415
Lys Thr Gly Thr Pro Ala Ala Gln Ile Gly Leu Lys Lys Gly Asp Val		

整理番号 156220

420                      425                      430  
 Ile Ile Gly Ala Asn Gln Gln Ala Val Lys Asn Ile Ala Glu Leu Arg  
 435                      440                      445  
 Lys Val Leu Asp Ser Lys Pro Ser Val Leu Ala Leu Asn Ile Gln Arg  
 450                      455                      460  
 Gly Asp Arg His Leu Pro Val Asn Ala Val Ile Ser Leu Asn Pro Phe  
 465                      470                      475                      480  
 Leu Lys Thr Gly Arg Gly Ser Pro Tyr Asn Leu  
                     485                      490

SEQ ID NO:15:

CTGGATGGGG AGGTGATTGG AGTG

24

SEQ ID NO:16:

GTCTCTGGGC CCCGGTTGTC TGTTG

25

SEQ ID NO:17:

Feature polymorphism at 1325

CCGGCCCTCG	CCCTGTCCGC	CGCCACCGCC	GCCGCCGCCA	GAGTCGCCAT	GCAGATCCCG	60
CGCGCCGCTC	TTCTCCCGCT	GCTGCTGCTG	CTGCTGGCGG	CGCCCGCCTC	GGCGCAGCTG	120
TCCCGGGCCG	GCGCTCGGC	GCCTTTGGCC	GCCGGGTGCC	CAGACCGCTG	CGAGCCGGCG	180
CGCTGCCCGC	CGCAGCCGGA	GCAC TGCGAG	GGCGGCCGGG	CCCGGGACGC	GTGCGGCTGC	240
TGCGAGGTGT	GCGCGCGGCC	CGAGGGCGCC	GCGTGCGGCC	TGCAGGAGGG	CCCGTGCGGC	300
GAGGGGCTGC	AGTGCGTGGT	GCCCTTCGGG	GTGCCAGCCT	CGGCCACGGT	GCGGCGGCGC	360
GCGCAGGCCG	GOCTCTGTGT	GTGCCCCAGC	AGCGAGCCGG	TGTGCGGCAG	CGACGCCAAC	420
ACCTACGCCA	ACCTGTGCCA	GCTGCGCGCC	GCCAGCCGCC	GCTCCGAGAG	GCTGCACCGG	480
CCGCCGGTCA	TGCTCTGCA	GCGCGGAGCC	TGCGGCCAAG	GGCAGGAAGA	TCCCAACAGT	540
TTGCGCCATA	AATATAACTT	TATCGCGGAC	GTGGTGGAGA	AGATCGCCCC	TGCCGTGGTT	600
CATATCGAAT	TGTTTCGCAA	GCTTCCGTTT	TCTAAACGAG	AGGTGCCGGT	GGCTAGTGGG	660
TCTGGGTTTA	TTGTGTGCGA	AGATGGACTG	ATCGTGACAA	ATGCCACAGT	GGTGACCAAC	720
AAGCACCGGG	TCAAAGTTGA	GCTGAAGAAC	GGTGCCACTT	ACGAAGCCAA	AATCAAGGAT	780
GTGGATGAGA	AAGCAGACAT	CGCACTCATC	AAAATTGACC	ACCAGGGCAA	GCTGOCTGTC	840
CTGTGTGCTT	GCCGCTCCTC	AGAGCTGCGG	CCGGGAGAGT	TCGTGGTCGC	CATCGGAAGC	900
CCGTTTTCCC	TTCAAAACAC	AGTCACCACC	GGGATCGTGR	GCACCACCCA	GCGAGGCGGC	960
AAAGAGCTGG	GGCTCCGCAA	CTCAGACATG	GACTACATCC	AGACCGACGC	CATCATCAAC	1020
TATGGAAACT	CGGGAGGCC	GTTAGTAAAC	CTGGACGGTG	AAGTGATTGG	AATTAACACT	1080
TTGAAAGTGA	CAGCTGGAAT	CTCCTTTGCA	ATCCCATCTG	ATAAGATTAA	AAAGTTCCTC	1140
ACGGAGTCCC	ATGACCGACA	GGCCAAAGGA	AAAGCCATCA	CCAAGAAGAA	GTATATTGGT	1200
ATCCGAATGA	TGTCACCTCAC	GTCCAGCAAA	GCCAAAGAGC	TGAAGGACCG	GCACCGGGAC	1260
TTCCAGACG	TGATCTCAGG	AGCGTATATA	ATTGAAGTAA	TTCTGATAC	CCCAGCAGAA	1320
GCTGKTGGTC	TCAAGGAAAA	CGACGTCATA	ATCAGCATCA	ATGGACAGTC	CGTGGTCTCC	1380
GCCAATGATG	TCAGCGACGT	CATTAAAGG	GAAAGCACCC	TGAACATGGT	GGTCCGCAGG	1440



整理番号 156220

GGTAATGAAG ATATCATGAT CACAGTGATT CCCGAAGAAA TTGACCCATA GGCAGAGGCA 1500  
 TGAGCTGGAC TTCATGTTTC CCTCAAAGAC TCTCCCGTGG ATGACGGATG AGGACTCTGG 1560  
 GCTGCTGGAA TAGGACACTC AAGACTTTTG ACTGCCATTT TGTGTTTCA GTGGAGACTC 1620  
 CCTGGCCAAC AGAATCCTTC TTGATAGTTT GCAGGCAAAA CAAATGTAAT GTTGCAGATC 1680  
 CGCAGGCAGA AGCTCTGCCC TTCTGTATCC TATGTATGCA GTGTGCTTTT TCTTGCCAGC 1740  
 TTGGGCCATT CTTGCTTAGA CAGTCAGCAT TTGTCTCTC CTTTAACTGA GTCATCATCT 1800  
 TAGTCCAACT AATGCAGTCG ATACAATGCG TAGATAGAAG AAGCCCCACG GGAGCCAGGA 1860  
 TGGGACTGGT CGTGTITGTG CTTTCTCCA AGTCAGCACC CAAAGGTCAA TGCACAGAGA 1920  
 CCCCCGGTGG GTGAGCGCTG GCTTCTCAA CGGCCGAAGT TGCCTCTTTT AGGAATCTCT 1980  
 TTGGAATTGG GAGCAGCATG ACTCTGAGTT TGAGCTATTA AAGTACTTCT TACAAA 2036

SEQ ID NO:18:

Feature - 213 Gly/val polymorph

Met Gln Ile Pro Arg Ala Ala Leu Leu Pro Leu Leu Leu Leu Leu Leu  
 1 5 10 15  
 Ala Ala Pro Ala Ser Ala Gln Leu Ser Arg Ala Gly Arg Ser Ala Pro  
 20 25 30  
 Leu Ala Ala Gly Cys Pro Asp Arg Cys Glu Pro Ala Arg Cys Pro Pro  
 35 40 45  
 Gln Pro Glu His Cys Glu Gly Gly Arg Ala Arg Asp Ala Cys Gly Cys  
 50 55 60  
 Cys Glu Val Cys Gly Ala Pro Glu Gly Ala Ala Cys Gly Leu Gln Glu  
 65 70 75 80  
 Gly Pro Cys Gly Glu Gly Leu Gln Cys Val Val Pro Phe Gly Val Pro  
 85 90 95  
 Ala Ser Ala Thr Val Arg Arg Arg Ala Gln Ala Gly Leu Cys Val Cys  
 100 105 110  
 Ala Ser Ser Glu Pro Val Cys Gly Ser Asp Ala Asn Thr Tyr Ala Asn  
 115 120 125  
 Leu Cys Gln Leu Arg Ala Ala Ser Arg Arg Ser Glu Arg Leu His Arg  
 130 135 140  
 Pro Pro Val Ile Val Leu Gln Arg Gly Ala Cys Gly Gln Gly Gln Glu  
 145 150 155 160  
 Asp Pro Asn Ser Leu Arg His Lys Tyr Asn Phe Ile Ala Asp Val Val  
 165 170 175  
 Glu Lys Ile Ala Pro Ala Val Val His Ile Glu Leu Phe Arg Lys Leu  
 180 185 190  
 Pro Phe Ser Lys Arg Glu Val Pro Val Ala Ser Gly Ser Gly Phe Ile  
 195 200 205  
 Val Ser Glu Asp Xaa Leu Ile Val Thr Asn Ala His Val Val Thr Asn  
 210 215 220  
 Lys His Arg Val Lys Val Glu Leu Lys Asn Gly Ala Thr Tyr Glu Ala

整理番号 156220

225                      230                      235                      240  
 Lys Ile Lys Asp Val Asp Glu Lys Ala Asp Ile Ala Leu Ile Lys Ile  
                          245                      250                      255  
 Asp His Gln Gly Lys Leu Pro Val Leu Leu Leu Gly Arg Ser Ser Glu  
                          260                      265                      270  
 Leu Arg Pro Gly Glu Phe Val Val Ala Ile Gly Ser Pro Phe Ser Leu  
                          275                      280                      285  
 Gln Asn Thr Val Thr Thr Gly Ile Val Ser Thr Thr Gln Arg Gly Gly  
                          290                      295                      300  
 Lys Glu Leu Gly Leu Arg Asn Ser Asp Met Asp Tyr Ile Gln Thr Asp  
 305                      310                      315                      320  
 Ala Ile Ile Asn Tyr Gly Asn Ser Gly Gly Pro Leu Val Asn Leu Asp  
                          325                      330                      335  
 Gly Glu Val Ile Gly Ile Asn Thr Leu Lys Val Thr Ala Gly Ile Ser  
                          340                      345                      350  
 Phe Ala Ile Pro Ser Asp Lys Ile Lys Lys Phe Leu Thr Glu Ser His  
                          355                      360                      365  
 Asp Arg Gln Ala Lys Gly Lys Ala Ile Thr Lys Lys Lys Tyr Ile Gly  
                          370                      375                      380  
 Ile Arg Met Met Ser Leu Thr Ser Ser Lys Ala Lys Glu Leu Lys Asp  
 385                      390                      395                      400  
 Arg His Arg Asp Phe Pro Asp Val Ile Ser Gly Ala Tyr Ile Ile Glu  
                          405                      410                      415  
 Val Ile Pro Asp Thr Pro Ala Glu Ala Gly Gly Leu Lys Glu Asn Asp  
                          420                      425                      430  
 Val Ile Ile Ser Ile Asn Gly Gln Ser Val Val Ser Ala Asn Asp Val  
                          435                      440                      445  
 Ser Asp Val Ile Lys Arg Glu Ser Thr Leu Asn Met Val Val Arg Arg  
                          450                      455                      460  
 Gly Asn Glu Asp Ile Met Ile Thr Val Ile Pro Glu Glu Ile Asp Pro  
 465                      470                      475                      480

SEQ ID NO:19:

ATGCTGAACA TCGGGAAAGC TTGGTTCTCG

30

SEQ ID NO:20:

CCACAGACA ACCGGGCCCA GAGACT

26

SEQ ID NO:21:

TGCCTCCTCG CCCGCCCTAC TCAGA

25

SEQ ID NO:22:

整理番号 156220

CGTGGATCCC GAGAAAGAGG CGCAGGACGA GGAGGCAGAA CCCGACTGGC GCGTAGAGCA 60  
 GCAGCACGAG CAGTAGGAAG CAGTCACCCG GAAGCCTGGG GGCAGAGAGC GAAGTGGTCA 120  
 GCGCCCGAAG GCCGAGAGCA CGCGGGGATC GGTCTCTTCC CGCCGGGTCT CTTACCGGTG 180  
 CGAGTCAAAG AGCCGCTCCG GCCCGGGCCC TGAGGGAAGC TCCATAACTG CTGCTTCAGG 240  
 AGCGCCCGGC CGTCGCCGCC GCCGCCATTT TCGCGCCCGG CCGCAGGGGC TCTTGGGAAG 300  
 GCGGAGTCTT TGGGCATCCG CCCGGGGTGA GGGGACCCGA AGTCCTGAGG CGCGCCGGAA 360  
 GGGCTAGCGG TCCCAGCATA CCCCGCGGCC CCTTGGGCGG TCTCACAACG CGCGTCCGGC 420  
 GGAGACCACA ATTCCCGCA TTCGTGGGGC AGGGAGGAGT CGGCCTCCCG GAATCCTGGT 480  
 CCCGGCGTGC ACTTCTGAAG GACTTCAGGT ACCGGCGTGC CCCGCGTCCT ACTGTCCGCC 540  
 TGCTCGCGTC CTGGGTGCCG CCTCTGAGTA GGGCGGGCGA GGAGGCA 587

SEQ ID NO:23:

CGTGGATCCC GAGAAAGAGG CGCAGGACGA GGAGGCAGAA CCCGACTGGC GCGTAGAGCA 60  
 GCAGCACGAG CAGTAGGAAG CAGTCACCCG GAAGCCTGGG GGCAGAGAGC GAAGTGGTCA 120  
 GCGCCCGAAG GCCGAGAGCA CGCGGGGATC GGTCTCTTCC CGCCGGGTCT CTTACCGGTG 180  
 CGAGTCAAAG AGCCGCTCCG GCCCGGGCCC TGAGGGAAGC TCCATAACTG CTGCTTCAGG 240  
 AGCGCCCGGC CGTCGCCGCC GCCGCCATTT TCGCGCCCGG CCGCAGGGGC TCTTGGGAAG 300  
 GCGGAGTCTT TGGGCATCCG CCCGGGGTGA GGGGACCCGA AGTCCTGAGG CGCGCCGGAA 360  
 GGGCTAGCGG TCCCAGCATA CCCCGCGGCC CCTTGGGCGG TCTCACAACG CGCGTCCGGC 420  
 GGAGACCACA ATTCCCGCA TTCGTGGGGC AGGGAGGAGT CGGCCTCCCG GAATCCTGGT 480  
 CCCGGCGTGC ACTTCTGAAG GACTTCAGGT ACCGGCGTGC CCCGCGTCCT ACTGTCCGCC 540  
 TGCTCGCGTC CTGGGTGCCG CCTCTGAGTA GGGCGGGCGA GGAGGCAGCC AAGGCGGAGC 600  
 TG ATG GCT GCG CCG AGG GCG GGG CGG GGT GCA GGC TGG AGC CTT CGG 647  
 Met Ala Ala Pro Arg Ala Gly Arg Gly Ala Gly Trp Ser Leu Arg  
 1 5 10 15

GCA TGG CGG GCT TTG GGG GGC ATT TGC TGG GGG AGG AGA CCC CGT TTG 695  
 Ala Trp Arg Ala Leu Gly Gly Ile Cys Trp Gly Arg Arg Pro Arg Leu  
 20 25 30

ACC CCT GAC CTC CGG GCC CTG CTG ACG TCA GGA ACT TCT GAC CCC CGG 743  
 Thr Pro Asp Leu Arg Ala Leu Leu Thr Ser Gly Thr Ser Asp Pro Arg  
 35 40 45

GCC CGA GTG ACT TAT GGG ACC CCC AGT CTC TGG GCC CGG TTG TCT GTT 791  
 Ala Arg Val Thr Tyr Gly Thr Pro Ser Leu Trp Ala Arg Leu Ser Val  
 50 55 60

GGG GTC ACT GAA CCC CGA GCA TGC CTG ACG TCT GGG ACC CCG GGT CCC 839  
 Gly Val Thr Glu Pro Arg Ala Cys Leu Thr Ser Gly Thr Pro Gly Pro  
 65 70 75

CGG GCA CAA CTG ACT GCG GTG ACC CCA GAT ACC AGG ACC CGG GAG GCC	887
Arg Ala Gln Leu Thr Ala Val Thr Pro Asp Thr Arg Thr Arg Glu Ala	
80 85 90 95	
TCA GAG AAC TCT GGA ACC CGT TCG CGC GCG TGG CTG GCG GTG GCG CTG	935
Ser Glu Asn Ser Gly Thr Arg Ser Arg Ala Trp Leu Ala Val Ala Leu	
100 105 110	
GGC GCT GGG GGG GCA GTG CTG TTG TTG TTG TGG GGC GGG GGT CGG GGT	983
Gly Ala Gly Gly Ala Val Leu Leu Leu Leu Trp Gly Gly Gly Arg Gly	
115 120 125	
CCT CCG GCC GTC CTC GCC GCC GTC CCT ACC CCG CCG CCC CCT TCT CCC	1031
Pro Pro Ala Val Leu Ala Ala Val Pro Ser Pro Pro Pro Ala Ser Pro	
130 135 140	
CGG AGT CAG TAC AAC TTC ATC GCA GAT GTG GTG GAG AAG ACA GCA CCT	1079
Arg Ser Gln Tyr Asn Phe Ile Ala Asp Val Val Glu Lys Thr Ala Pro	
145 150 155	
GCC GTG GTC TAT ATC GAG ATC CTG GAC CGG CAC CCT TTC TTG GGC CGC	1127
Ala Val Val Tyr Ile Glu Ile Leu Asp Arg His Pro Phe Leu Gly Arg	
160 165 170 175	
GAG GTC CCT ATC TCG AAC GGC TCA GGA TTC GTG GTG GCT GCC GAT GGG	1175
Glu Val Pro Ile Ser Asn Gly Ser Gly Phe Val Val Ala Ala Asp Gly	
180 185 190	
CTC ATT GTC ACC AAC GCC CAT GTG GTG GCT GAT CGG CGC AGA GTC CGT	1223
Leu Ile Val Thr Asn Ala His Val Val Ala Asp Arg Arg Arg Val Arg	
195 200 205	
GTG AGA CTG CTA AGC GGC GAC ACG TAT GAG GCC GTG GTC ACA GCT GTG	1271
Val Arg Leu Leu Ser Gly Asp Thr Tyr Glu Ala Val Val Thr Ala Val	
210 215 220	
GAT CCC GTG GCA GAC ATC GCA ACG CTG AGG ATT CAG ACT AAG GAG CCT	1319
Asp Pro Val Ala Asp Ile Ala Thr Leu Arg Ile Gln Thr Lys Glu Pro	
225 230 235	
CTC CCC ACG CTG CCT CTG GGA CGC TCA GCT GAT GTC CCG CAA GGG GAG	1367
Leu Pro Thr Leu Pro Leu Gly Arg Ser Ala Asp Val Arg Gln Gly Glu	

整理番号 156220

240	245	250	255	
TTT GTT GTT GCC ATG GGA AGT CCC TTT GCA CTG CAG AAC ACG ATC ACA				1415
Phe Val Val Ala Met Gly Ser Pro Phe Ala Leu Gln Asn Thr Ile Thr				
260	265	270		
TCC GGC ATT GTT AGC TCT GCT CAG CGT CCA GCC AGA GAC CTG GGA CTC				1463
Ser Gly Ile Val Ser Ser Ala Gln Arg Pro Ala Arg Asp Leu Gly Leu				
275	280	285		
CCC CAA ACC AAT GTG GAA TAC ATT CAA ACT GAT GCA GCT ATT GAT TTT				1511
Pro Gln Thr Asn Val Glu Tyr Ile Gln Thr Asp Ala Ala Ile Asp Phe				
290	295	300		
GGA AAC TCT GGA GGT CCC CTG GTT AAC CTG GAT GGG GAG GTG ATT GGA				1559
Gly Asn Ser Gly Gly Pro Leu Val Asn Leu Asp Gly Glu Val Ile Gly				
305	310	315		
GTG AAC ACC ATG AAG GTC ACA GCT GGA ATC TCC TTT GCC ATC CCT TCT				1607
Val Asn Thr Met Lys Val Thr Ala Gly Ile Ser Phe Ala Ile Pro Ser				
320	325	330	335	
GAT CGT CTT CGA GAG TTT CTG CAT CGT GGG GAA AAG AAG AAT TCC TCC				1655
Asp Arg Leu Arg Glu Phe Leu His Arg Gly Glu Lys Lys Asn Ser Ser				
340	345	350		
TCC GGA ATC AGT GGG TCC CAG CGG CGC TAC ATT GGG GTG ATG ATG CTG				1703
Ser Gly Ile Ser Gly Ser Gln Arg Arg Tyr Ile Gly Val Met Met Leu				
355	360	365		
ACC CTG AGT CCC AGC ATC CTT GCT GAA CTA CAG CTT CGA GAA CCA ACC				1751
Thr Leu Ser Pro Ser Ile Leu Ala Glu Leu Gln Leu Arg Glu Pro Ser				
370	375	380		
TTT CCC GAT GTT CAG CAT GGT GTA CTC ATC CAT AAA GTC ATC CTG GGC				1799
Phe Pro Asp Val Gln His Gly Val Leu Ile His Lys Val Ile Leu Gly				
385	390	395		
TCC CCT GCA CAC CGG GCT GGT CTG CGG CCT GGT GAT GTG ATT TTG GCC				1847
Ser Pro Ala His Arg Ala Gly Leu Arg Pro Gly Asp Val Ile Leu Ala				
400	405	410	415	

整理番号 156220

ATT GGG GAG CAG ATG GTA CAA AAT GCT GAA GAT GTT TAT GAA GCT GTT 1895  
 Ile Gly Glu Gln Met Val Gln Asn Ala Glu Asp Val Tyr Glu Ala Val  
 420 425 430

CGA ACC CAA TCC CAG TTG GCA GTG CAG ATC CGG CGG GGA CGA GAA ACA 1943  
 Arg Thr Gln Ser Gln Leu Ala Val Gln Ile Arg Arg Gly Arg Glu Thr  
 435 440 445

CTG ACC TTA TAT GTG ACC CCT GAG GTC ACA GAA TGAATAGATC ACCAAGAGTA 1996  
 Leu Thr Leu Tyr Val Thr Pro Glu Val Thr Glu  
 450 455

TGAGGCTCCT GCTCTGATT CTTCTTGCC TTTCTGGCTG AGGTTCTGAG GGCACCGAGA 2056  
 CAGAGGGTTA AATGAACCAG TGGGGGCAGG TCCCTCCAC CACCAGCACT GACTCCTGGG 2116  
 CTCTGAAGAA TCACAGAAAC ACTTTTATA TAAATATAAA TTATACCTAG CAACATAAAA 2176  
 AAAAAAAAAA A 2187

SEQ ID NO:24:

CGTGGATCCC GAGAAAGAGG CGCAGGACGA GGAGGCAGAA CCCGACTGGC GCGTAGAGCA 60  
 GCAGCAGGAG CAGTAGGAAG CAGTCACCCG GAAGCCTGGG GGCAGAGGCC GAAGTGGTCA 120  
 GCGCGCGAAG GCGAGAGCA CGCGGGGATC GGTCTCTTCC CGCGGGGTCT CTTACCGGTG 180  
 CGAGTCAAAG AGCGCTCCG GCCCGGGCCC TGAGGGAAGC TCCATAACTG CTGCTTCAGG 240  
 AGCGCCCGGC CGTCCCGCC GCCGCCATTT TCGCGCTCGG CGCAGGGGC TCTTGGGAG 300  
 GCGGAGTCTT TGGGCATCCG CCCGGGGTGA GGGGACCCGA AGTCCTGAGG CGCGCCGGAA 360  
 GGGCTAGCGG TCCAGCATA CCCCGCGGCC CCTTGGGCGG TCTCACAAC CGCGTCCGGC 420  
 GGAGACCACA ATTCCCGGCA TTCGTGGGGC AGGGAGGAGT CGGCCTCCCG GAATCCTGGT 480  
 CCCGGCGTGC ACTTCTGAAG GACTTCAGGT ACCGGCGTGC CCCGCGTCTT ACTGTCCGCC 540  
 TGCTCGCGTC CTGGGTGCCG CCTCTGAGTA GGGCGGGCGA GGAGGCAGCC AAGGCGGAGC 600  
 TG ATG GCT GCG CCG AGG GCG GGG CGG GGT GCA GGC TGG AGC CTT CGG 647  
 Met Ala Ala Pro Arg Ala Gly Arg Gly Ala Gly Trp Ser Leu Arg  
 1 5 10 15

GCA TGG CGG GCT TTG GGG GGC ATT CGC TGG GGG AGG AGA CCC CGT TTG 695  
 Ala Trp Arg Ala Leu Gly Gly Ile Arg Trp Gly Arg Arg Pro Arg Leu  
 20 25 30

ACC CCT GAC CTC CGG GCC CTG CTG ACG TCA GGA ACT TCT GAC CCC CGG 743  
 Thr Pro Asp Leu Arg Ala Leu Leu Thr Ser Gly Thr Ser Asp Pro Arg  
 35 40 45

GCC CGA GTG ACT TAT GGG ACC CCC AGT CTC TGG GCC CGG TTG TCT GTT 791  
 Ala Arg Val Thr Tyr Gly Thr Pro Ser Leu Trp Ala Arg Leu Ser Val

整理番号 156220

50	55	60	
GGG GTC ACT GAA CCC CGA GCA TGC CTG ACG TCT GGG ACC CCG GGT CCC			839
Gly Val Thr Glu Pro Arg Ala Cys Leu Thr Ser Gly Thr Pro Gly Pro			
65	70	75	
CGG GCA CAA CTG ACT GCG GTG ACC CCA GAT ACC AGG ACC CGG GAG GCC			887
Arg Ala Gln Leu Thr Ala Val Thr Pro Asp Thr Arg Thr Arg Glu Ala			
80	85	90	95
TCA GAG AAC TCT GGA ACC CGT TCG CGC GCG TGG CTG GCG GTG GCG CTG			935
Ser Glu Asn Ser Gly Thr Arg Ser Arg Ala Trp Leu Ala Val Ala Leu			
100	105	110	
GGC GCT GGG CGG GCA GTG CTG TTG TTG TTG TGG GGC GGG GGT CGG GGT			983
Gly Ala Gly Gly Ala Val Leu Leu Leu Trp Gly Gly Gly Arg Gly			
115	120	125	
CCT CCG GCC GTC CTC GCC GCC GTC CCT AGC CCG CCG CCC GCT TCT CCC			1031
Pro Pro Ala Val Leu Ala Ala Val Pro Ser Pro Pro Pro Ala Ser Pro			
130	135	140	
CGG AGT CAG TAC AAC TTC ATC GCA GAT GTG GTG GAG AAG ACA GCA CCT			1079
Arg Ser Gln Tyr Asn Phe Ile Ala Asp Val Val Glu Lys Thr Ala Pro			
145	150	155	
GCC GTG GTC TAT ATC GAG ATC CTG GAC CGG CAC CCT TTC TTG GGC CGC			1127
Ala Val Val Tyr Ile Glu Ile Leu Asp Arg His Pro Phe Leu Gly Arg			
160	165	170	175
GAG GTC CCT ATC TCG AAC GGC TCA GGA TTC GTG GTG GCT GCC GAT GGG			1175
Glu Val Pro Ile Ser Asn Gly Ser Gly Phe Val Val Ala Ala Asp Gly			
180	185	190	
CTC ATT GTC ACC AAC GCC CAT GTG GTG GCT GAT CCG CGC AGA GTC CGT			1223
Leu Ile Val Thr Asn Ala His Val Val Ala Asp Arg Arg Arg Val Arg			
195	200	205	
GTG AGA CTG CTA AGC GGC GAC ACG TAT GAG GCC GTG GTC ACA GCT GTG			1271
Val Arg Leu Leu Ser Gly Asp Thr Tyr Glu Ala Val Val Thr Ala Val			
210	215	220	

整理番号 156220

GAT CCC GTG GCA GAC ATC GCA ACG CTG AGG ATT CAG ACT AAG GAG CCT Asp Pro Val Ala Asp Ile Ala Thr Leu Arg Ile Gln Thr Lys Glu Pro 225 230 235	1319
CTC CCC ACG CTG CCT CTG GGA CGC TCA GCT GAT GTC CGG CAA GGG GAG Leu Pro Thr Leu Pro Leu Gly Arg Ser Ala Asp Val Arg Gln Gly Glu 240 245 250 255	1367
TTT GTT GTT GCC ATG GGA AGT CCC TTT GCA CTG CAG AAC ACG ATC ACA Phe Val Val Ala Met Gly Ser Pro Phe Ala Leu Gln Asn Thr Ile Thr 260 265 270	1415
TCC GGC ATT GTT AGC TCT GCT CAG CGT CCA GCC AGA GAC CTG GGA CTC Ser Gly Ile Val Ser Ser Ala Gln Arg Pro Ala Arg Asp Leu Gly Leu 275 280 285	1463
CCC CAA ACC AAT GTG GAA TAC ATT CAA ACT GAT GCA GCT ATT GAT TTT Pro Gln Thr Asn Val Glu Tyr Ile Gln Thr Asp Ala Ala Ile Asp Phe 290 295 300	1511
GGA AAC TCT GGA GGT CCC CTG GTT AAC CTG GAT GGG GAG GTG ATT GGA Gly Asn Ser Gly Gly Pro Leu Val Asn Leu Asp Gly Glu Val Ile Gly 305 310 315	1559
GTG AAC ACC ATG AAG GTC ACA GCT GGA ATC TCC TTT GCC ATC CCT TCT Val Asn Thr Met Lys Val Thr Ala Gly Ile Ser Phe Ala Ile Pro Ser 320 325 330 335	1607
GAT CGT CTT CGA GAG TTT CTG CAT CGT GGG GAA AAG AAG AAT TCC TCC Asp Arg Leu Arg Glu Phe Leu His Arg Gly Glu Lys Lys Asn Ser Ser 340 345 350	1655
TCC GGA ATC AGT GGG TCC CAG CGG CGC TAC ATT GGG GTG ATG ATG CTG Ser Gly Ile Ser Gly Ser Gln Arg Arg Tyr Ile Gly Val Met Met Leu 355 360 365	1703
ACC CTG AGT CCC AGC ATC CTT GCT GAA CTA CAG CTT CGA GAA CCA AGC Thr Leu Ser Pro Ser Ile Leu Ala Glu Leu Gln Leu Arg Glu Pro Ser 370 375 380	1751
TTT CCC GAT GTT CAG CAT GGT GTA CTC ATC CAT AAA GTC ATC CTG GGC Phe Pro Asp Val Gln His Gly Val Leu Ile His Lys Val Ile Leu Gly	1799



整理番号 156220

385	390	395	
TCC CCT GCA CAC CGG GCT GGT CTG CGG CCT GGT GAT GTG ATT TTG GCC			1847
Ser Pro Ala His Arg Ala Gly Leu Arg Pro Gly Asp Val Ile Leu Ala			
400	405	410	415
ATT GGG GAG CAG ATG GTA CAA AAT GCT GAA GAT GTT TAT GAA GCT GTT			1895
Ile Gly Glu Gln Met Val Gln Asn Ala Glu Asp Val Tyr Glu Ala Val			
	420	425	430
CGA ACC CAA TCC CAG TTG GCA GTG CAG ATC CGG CGG GGA CGA GAA ACA			1943
Arg Thr Gln Ser Gln Leu Ala Val Gln Ile Arg Arg Gly Arg Glu Thr			
	435	440	445
CTG ACC TTA TAT GTG ACC CCT GAG GTC ACA GAA TGAATAGATC ACCAAGAGTA			1996
Leu Thr Leu Tyr Val Thr Pro Glu Val Thr Glu			
	450	455	
TGAGGCTCCT GCTCTGATTI CCTCCTIGCC TITCTGGCTG AGGTICTGAG GGCACCGAGA			2056
CAGAGGGTTA AATGAACCAG TGGGGGCAGG TCCCTCCAAC CACCAGCACT GACTCCTGGG			2116
CTCTGAAGAA TCACAGAAAC ACTTTTATA TAAATAAATA TTATACCTAG CACATATAAA			2176
AAAAAAAAA A			2187

SEQ ID NO:25:

Met Ala Ala Pro Arg Ala Gly Arg Gly Ala Gly Trp Ser Leu Arg Ala			
1	5	10	15
Trp Arg Ala Leu Gly Gly Ile Arg Trp Gly Arg Arg Pro Arg Leu Thr			
	20	25	30
Pro Asp Leu Arg Ala Leu Leu Thr Ser Gly Thr Ser Asp Pro Arg Ala			
	35	40	45
Arg Val Thr Tyr Gly Thr Pro Ser Leu Trp Ala Arg Leu Ser Val Gly			
	50	55	60
Val Thr Glu Pro Arg Ala Cys Leu Thr Ser Gly Thr Pro Gly Pro Arg			
	65	70	75
Ala Gln Leu Thr Ala Val Thr Pro Asp Thr Arg Thr Arg Glu Ala Ser			
	85	90	95
Glu Asn Ser Gly Thr Arg Ser Arg Ala Trp Leu Ala Val Ala Leu Gly			
	100	105	110
Ala Gly Gly Ala Val Leu Leu Leu Leu Trp Gly Gly Gly Arg Gly Pro			
	115	120	125
Pro Ala Val Leu Ala Ala Val Pro Ser Pro Pro Pro Ala Ser Pro Arg			
	130	135	140

整理番号 156220

```

Ser Gln Tyr Asn Phe Ile Ala Asp Val Val Glu Lys Thr Ala Pro Ala
145                      150                      155                      160
Val Val Tyr Ile Glu Ile Leu Asp Arg His Pro Phe Leu Gly Arg Glu
                      165                      170                      175
Val Pro Ile Ser Asn Gly Ser Gly Phe Val Val Ala Ala Asp Gly Leu
                      180                      185                      190
Ile Val Thr Asn Ala His Val Val Ala Asp Arg Arg Arg Val Arg Val
                      195                      200                      205
Arg Leu Leu Ser Gly Asp Thr Tyr Glu Ala Val Val Thr Ala Val Asp
                      210                      215                      220
Pro Val Ala Asp Ile Ala Thr Leu Arg Ile Gln Thr Lys Glu Pro Leu
225                      230                      235                      240
Pro Thr Leu Pro Leu Gly Arg Ser Ala Asp Val Arg Gln Gly Glu Phe
                      245                      250                      255
Val Val Ala Met Gly Ser Pro Phe Ala Leu Gln Asn Thr Ile Thr Ser
                      260                      265                      270
Gly Ile Val Ser Ser Ala Gln Arg Pro Ala Arg Asp Leu Gly Leu Pro
                      275                      280                      285
Gln Thr Asn Val Glu Tyr Ile Gln Thr Asp Ala Ala Ile Asp Phe Gly
                      290                      295                      300
Asn Ser Gly Gly Pro Leu Val Asn Leu Asp Gly Glu Val Ile Gly Val
305                      310                      315                      320
Asn Thr Met Lys Val Thr Ala Gly Ile Ser Phe Ala Ile Pro Ser Asp
                      325                      330                      335
Arg Leu Arg Glu Phe Leu His Arg Gly Glu Lys Lys Asn Ser Ser Ser
                      340                      345                      350
Gly Ile Ser Gly Ser Gln Arg Arg Tyr Ile Gly Val Met Met Leu Thr
                      355                      360                      365
Leu Ser Pro Ser Ile Leu Ala Glu Leu Gln Leu Arg Glu Pro Ser Phe
                      370                      375                      380
Pro Asp Val Gln His Gly Val Leu Ile His Lys Val Ile Leu Gly Ser
385                      390                      395                      400
Pro Ala His Arg Ala Gly Leu Arg Pro Gly Asp Val Ile Leu Ala Ile
                      405                      410                      415
Gly Glu Gln Met Val Gln Asn Ala Glu Asp Val Tyr Glu Ala Val Arg
                      420                      425                      430
Thr Gln Ser Gln Leu Ala Val Gln Ile Arg Arg Gly Arg Glu Thr Leu
                      435                      440                      445
Thr Leu Tyr Val Thr Pro Glu Val Thr Glu
                      450                      455

```

SEQ ID NO:26:

整理番号 156220

CGTGGATCCC GAGAAAGAGG CGCAGGACGA GGAGGCAGAA CCCGACTGGC GCGTAGAGCA	60
GCAGCACGAG CAGTAGGAAG CAGTCACCCG GAAGCCTGGG GGCGAGAGGC GAAGTGGTCA	120
GGCGCOGAAG GCCGAGAGCA CGCGGGGATC GGTCTCTTCC CGCGGGTCT CTTACCGGTG	180
CGAGTCAAAG AGCCGCTCOG GCCCGGGCCC TGAGGGAAGC TCCATAACTG CTGCTTCAGG	240
AGCGOCCGGC CGTCGCCGCC GCCGCCATTT TCGCGCCCGG CCGCAGGGGC TCTTGGGAAG	300
GCGGAGTCTT TGGGCATCOG CCGGGGGTGA GGGGACCGA AGTCCTGAGG CGCGCCGGAA	360
GGECTAGCGG TCCAGCATA CCCCGCGGCC CCTTGGGCGG TCTCACAAC TCCGTCGGCC	420
GGAGAACACA ATTCGCCGCA TTCGTGGGGC AGGGAGGAGT CGGCCTCCCG GAATCCTGGT	480
CCCGGCGTGC ACTTCTGAAG GACTTCAGGT ACCGGCGTGC CCGCGTCTCT ACTGTCCGCC	540
TGCTCCGCTC CTGGGTGCCG CCTCTGAGTA GGGCGGGCGA GGAGGCAGCC AAGGCGGAGC	600
TA ATG GCT GCG CCG AGG GCG GGG CGG GGT GCA GGC TGG AGC CTT CGG	647
Met Ala Ala Pro Arg Ala Gly Arg Gly Ala Gly Trp Ser Leu Arg	
1 5 10 15	
GCA TGG CGG GCT TTG GGG GGC ATT CGC TGG GGG AGG AGA CCC CGT TTG	695
Ala Trp Arg Ala Leu Gly Gly Ile Arg Trp Gly Arg Arg Pro Arg Leu	
20 25 30	
ACC CCT GAC CTC CGG GCC CTG CTG ACG TCA GGA ACT TCT GAC CCC CGG	743
Thr Pro Asp Leu Arg Ala Leu Leu Thr Ser Gly Thr Ser Asp Pro Arg	
35 40 45	
GCC CGA GTG ACT TAT GGG ACC CCC AGT CTC TGG GCC CGG TTG TCT GTT	791
Ala Arg Val Thr Tyr Gly Thr Pro Ser Leu Trp Ala Arg Leu Ser Val	
50 55 60	
GGG GTC ACT GAA CCC CGA GCA TGC CTG ACG TCT GGG ACC CCG GGT CCC	839
Gly Val Thr Glu Pro Arg Ala Cys Leu Thr Ser Gly Thr Pro Gly Pro	
65 70 75	
CGG GCA CAA CTG ACT GCG GTG ACC CCA GAT ACC AGG ACC CGG GAG GCC	887
Arg Ala Gln Leu Thr Ala Val Thr Pro Asp Thr Arg Thr Arg Glu Ala	
80 85 90 95	
TCA GAG AAC TCT GGA ACC CGT TCG CGC GCG TGG CTG GCG GTG GCG CTG	935
Ser Glu Asn Ser Gly Thr Arg Ser Arg Ala Trp Leu Ala Val Ala Leu	
100 105 110	
GGC GCT GGG GGG GCA GTG CTG TTG TTG TTG TGG GGC GGG GGT CGG GGT	983
Gly Ala Gly Gly Ala Val Leu Leu Leu Leu Trp Gly Gly Gly Arg Gly	
115 120 125	

整理番号 156220

CCT CCG GCC GTC CTC GCC GCC GTC OCT AGC CCG CCG CCC GCT TCT CCC Pro Pro Ala Val Leu Ala Ala Val Pro Ser Pro Pro Pro Ala Ser Pro 130 135 140	1031
CGG AGT CAG TAC AAC TTC ATC GCA GAT GTG GTG GAG AAG ACA GCA CCT Arg Ser Gln Tyr Asn Phe Ile Ala Asp Val Val Glu Lys Thr Ala Pro 145 150 155	1079
GCC GTG GTC TAT ATC GAG ATC CTG GAC CCG CAC CCT TTC TTG GGC CGC Ala Val Val Tyr Ile Glu Ile Leu Asp Arg His Pro Phe Leu Gly Arg 160 165 170 175	1127
GAG GTC CCT ATC TCG AAC GGC TCA GGA TTC GTG GTG GCT GCC GAT CCG Glu Val Pro Ile Ser Asn Gly Ser Gly Phe Val Val Ala Ala Asp Gly 180 185 190	1175
CTC ATT GTC ACC AAC GGC CAT GTG GTG GCT GAT CCG CGC AGA GTC CGT Leu Ile Val Thr Asn Ala His Val Val Ala Asp Arg Arg Arg Val Arg 195 200 205	1223
GTG AGA CTG CTA AGC GGC GAC ACG TAT GAG GCC GTG GTC ACA GCT GTG Val Arg Leu Leu Ser Gly Asp Thr Tyr Glu Ala Val Val Thr Ala Val 210 215 220	1271
GAT CCC GTG GCA GAC ATC GCA ACG CTG AGG ATT CAG ACT AAG GAG CCT Asp Pro Val Ala Asp Ile Ala Thr Leu Arg Ile Gln Thr Lys Glu Pro 225 230 235	1319
CTC CCC ACG CTG CCT CTG GGA CGC TCA GCT GAT GTC CGG CAA GGG GAG Leu Pro Thr Leu Pro Leu Gly Arg Ser Ala Asp Val Arg Gln Gly Glu 240 245 250 255	1367
TTT GTT GTT GCC ATG GGA AGT CCC TTT GCA CTG CAG AAC ACG ATC ACA Phe Val Val Ala Met Gly Ser Pro Phe Ala Leu Gln Asn Thr Ile Thr 260 265 270	1415
TCC GGC ATT GTT AGC TCT GCT CAG CGT CCA GCC AGA GAC CTG GGA CTC Ser Gly Ile Val Ser Ser Ala Gln Arg Pro Ala Arg Asp Leu Gly Leu 275 280 285	1463
CCC CAA ACC AAT GTG GAA TAC ATT CAA ACT GAT GCA GCT ATT GAT TTT Pro Gln Thr Asn Val Glu Tyr Ile Gln Thr Asp Ala Ala Ile Asp Phe	1511

58

整理番号 156220

Pro Asp Leu Arg Ala Leu Leu Thr Ser Gly Thr Ser Asp Pro Arg Ala  
 35 40 45  
 Arg Val Thr Tyr Gly Thr Pro Ser Leu Trp Ala Arg Leu Ser Val Gly  
 50 55 60  
 Val Thr Glu Pro Arg Ala Cys Leu Thr Ser Gly Thr Pro Gly Pro Arg  
 65 70 75 80  
 Ala Gln Leu Thr Ala Val Thr Pro Asp Thr Arg Thr Arg Glu Ala Ser  
 85 90 95  
 Glu Asn Ser Gly Thr Arg Ser Arg Ala Trp Leu Ala Val Ala Leu Gly  
 100 105 110  
 Ala Gly Gly Ala Val Leu Leu Leu Leu Trp Gly Gly Gly Arg Gly Pro  
 115 120 125  
 Pro Ala Val Leu Ala Ala Val Pro Ser Pro Pro Pro Ala Ser Pro Arg  
 130 135 140  
 Ser Gln Tyr Asn Phe Ile Ala Asp Val Val Glu Lys Thr Ala Pro Ala  
 145 150 155 160  
 Val Val Tyr Ile Glu Ile Leu Asp Arg His Pro Phe Leu Gly Arg Glu  
 165 170 175  
 Val Pro Ile Ser Asn Gly Ser Gly Phe Val Val Ala Ala Asp Gly Leu  
 180 185 190  
 Ile Val Thr Asn Ala His Val Val Ala Asp Arg Arg Arg Val Arg Val  
 195 200 205  
 Arg Leu Leu Ser Gly Asp Thr Tyr Glu Ala Val Val Thr Ala Val Asp  
 210 215 220  
 Pro Val Ala Asp Ile Ala Thr Leu Arg Ile Gln Thr Lys Glu Pro Leu  
 225 230 235 240  
 Pro Thr Leu Pro Leu Gly Arg Ser Ala Asp Val Arg Gln Gly Glu Phe  
 245 250 255  
 Val Val Ala Met Gly Ser Pro Phe Ala Leu Gln Asn Thr Ile Thr Ser  
 260 265 270  
 Gly Ile Val Ser Ser Ala Gln Arg Pro Ala Arg Asp Leu Gly Leu Pro  
 275 280 285  
 Gln Thr Asn Val Glu Tyr Ile Gln Thr Asp Ala Ala Ile Asp Phe Gly  
 290 295 300  
 Asn Ser Gly Gly Pro Leu Val Asn Leu Val Ser Glu Thr Ser Phe Leu  
 305 310 315 320  
 Pro Arg Ile Pro Ala Pro Gly Gln Cys Gly Lys Gly Arg Phe Pro Leu  
 325 330 335  
 Ile Gln Gly Cys Leu Val Lys Phe Leu Ser Ser Ser Leu Leu Ala Ile  
 340 345 350  
 Ser Gln Tyr Pro Thr Arg Ser Pro Gln His Leu Leu Val Leu Leu Phe  
 355 360 365

整理番号 156220

Gly Cys Pro His Pro Leu Leu Phe Val  
370 375

SEQ ID NO:28:

CGTGGATCCC GAGAAAGAGG CGCAGGACGA GGAGGCAGAA CCGACTGGC GCGTAGAGCA	60
GCAGCAOGAG CACTAGGAAG CAGTCACCCG GAAGCCTGGG GCGAGAGGC GAAGTGGTCA	120
GGCGCOGAAG GCCGAGAGCA CGCGGGGATC GGTCTCTTCC CGCCGGGTCT CTTACCGGTG	180
CGAGTCAAAG AGCOGCTCCG GCGCGGGCCC TGAGGGAAGC TCCATAACTG CTGCTTCAGG	240
AGCGCCOOGG CGTOGCCGCC GCGGCCATTT TCGCGCCCGG CCGCAGGGGC TCTTGGGAAG	300
GCGGAGTCTT TGGGCATCCG CCCGGGGTGA GGGGACCCGA AGTCCTGAGG CGCGCCGGAA	360
GGGCTIAGCGG TCCCAGCATA CCCCGCGGCC CCTTGGGCGG TCTCACAAC TCGCTCCGGC	420
GGAGACCACA ATTCCCGCA TTCGTGGGGC AGGGAGGAGT CGGCCTCCCG GAATCCTGGT	480
CCCGGCGTGC ACTTCTGAAG GACTTCAGGT ACCGCGGTGC CCCGCGTCTT ACTGTCCGCC	540
TGCTCGCGTC CTGGGTGCGG CCTCTGAGTA GGGCGGGCGA GGAGGCAGCC AAGGCGGAGC	600
TG ATG GCT GCG CCG AGG GCG GGG CCG GGT GCA GGC TGG AGC CTT CGG	647
Met Ala Ala Pro Arg Ala Gly Arg Gly Ala Gly Trp Ser Leu Arg	
1 5 10 15	
GCA TGG CCG GCT TTG GGG GGC ATT CGC TGG GGG AGG AGA CCC CGT TTG	695
Ala Trp Arg Ala Leu Gly Gly Ile Arg Trp Gly Arg Arg Pro Arg Leu	
20 25 30	
ACC CCT GAC CTC CGG GCC CTG CTG ACG TCA GGA ACT TCT GAC CCC CGG	743
Thr Pro Asp Leu Arg Ala Leu Leu Thr Ser Gly Thr Ser Asp Pro Arg	
35 40 45	
GCC CGA GTG ACT TAT GGG ACC CCC AGT CTC TGG GCC CGG TTG TCT GTT	791
Ala Arg Val Thr Tyr Gly Thr Pro Ser Leu Trp Ala Arg Leu Ser Val	
50 55 60	
GGG GTC ACT GAA CCC CGA GCA TGC CTG ACG TCT GGG ACC CCG GGT CCC	839
Gly Val Thr Glu Pro Arg Ala Cys Leu Thr Ser Gly Thr Pro Gly Pro	
65 70 75	
CGG GCA CAA CTG ACT GCG GTG ACC CCA GAT ACC AGG ACC CGG GAG GCC	887
Arg Ala Gln Leu Thr Ala Val Thr Pro Asp Thr Arg Thr Arg Glu Ala	
80 85 90 95	
TCA GAG AAC TCT GGA ACC CGT TCG CGC GCG TGG CTG GCG GTG GCG CTG	935
Ser Glu Asn Ser Gly Thr Arg Ser Arg Ala Trp Leu Ala Val Ala Leu	
100 105 110	

整理番号 156220

GGC GCT GGG GGG GCA GTG CTG TTG TTG TTG TGG GGC GGG GGT CGG GGT Gly Ala Gly Gly Ala Val Leu Leu Leu Leu Trp Gly Gly Gly Arg Gly 115 120 125	983
CCT CCG GCC GTC CTC GCC GCC GTC CCT AGC CCG CCG CCC GCT TCT CCC Pro Pro Ala Val Leu Ala Ala Val Pro Ser Pro Pro Pro Ala Ser Pro 130 135 140	1031
CGG AGT CAG TAC AAC TTC ATC GCA GAT GTG GTG GAG AAG ACA GCA CCT Arg Ser Gln Tyr Asn Phe Ile Ala Asp Val Val Glu Lys Thr Ala Pro 145 150 155	1079
GCC GTG GTC TAT ATC GAG ATC CTG GAC CGG CAC CCT TTC TTG GGC CGC Ala Val Val Tyr Ile Glu Ile Leu Asp Arg His Pro Phe Leu Gly Arg 160 165 170 175	1127
GAG GTC CCT ATC TCG AAC GGC TCA GGA TTC GTG GTG GCT GCC GAT GGG Glu Val Pro Ile Ser Asn Gly Ser Gly Phe Val Val Ala Ala Asp Gly 180 185 190	1175
CTC ATT GTC ACC AAC GCC CAT GTG GTG GCT GAT CGG CGC AGA GTC CGT Leu Ile Val Thr Asn Ala His Val Val Ala Asp Arg Arg Arg Val Arg 195 200 205	1223
GTG AGA CTG CTA AGC GGC GAC ACG TAT GAG GCC GTG GTC ACA GCT GTG Val Arg Leu Leu Ser Gly Asp Thr Tyr Glu Ala Val Val Thr Ala Val 210 215 220	1271
GAT CCC GTG GCA GAC ATC GCA ACG CTG AGG ATT CAG ACT AAG GAG CCT Asp Pro Val Ala Asp Ile Ala Thr Leu Arg Ile Gln Thr Lys Glu Pro 225 230 235	1319
CTC CCC ACG CTG CCT CTG GGA CGC TCA GCT GAT GTC CGG CAA GGG GAG Leu Pro Thr Leu Pro Leu Gly Arg Ser Ala Asp Val Arg Gln Gly Glu 240 245 250 255	1367
TTT GTT GTT GCC ATG GGA AGT CCC TTT GCA CTG CAG AAC ACG ATC ACA Phe Val Val Ala Met Gly Ser Pro Phe Ala Leu Gln Asn Thr Ile Thr 260 265 270	1415
TCC GGC ATT GTT AGC TCT GCT CAG CGT CCA GCC AGA GAC CTG GGA CTC Ser Gly Ile Val Ser Ser Ala Gln Arg Pro Ala Arg Asp Leu Gly Leu	1463



生理番号 156220

275	280	285	
CCC CAA ACC AAT GTG GAA TAC ATT CAA ACT GAT GCA GCT ATT GAT TTT			1511
Pro Gln Thr Asn Val Glu Tyr Ile Gln Thr Asp Ala Ala Ile Asp Phe			
290	295	300	
GGA AAC TCT GGA GGT CCC CTG GTT AAC CTG GCT AGG GAA CTG GGG GCT			1559
Gly Asn Ser Gly Gly Pro Leu Val Asn Leu Ala Arg Glu Leu Gly Ala			
305	310	315	
GTA TCC CTG CAG GAT GGG GAG GTG ATT GGA GTG AAC ACC ATG AAG GTC			1607
Val Ser Leu Gln Asp Gly Glu Val Ile Gly Val Asn Thr Met Lys Val			
320	325	330	335
ACA GCT GGA ATC TCC TTT GCC ATC CCT TCT GAT CGT CTT CGA GAG TTT			1655
Thr Ala Gly Ile Ser Phe Ala Ile Pro Ser Asp Arg Leu Arg Glu Phe			
340	345	350	
CTG CAT CGT GGG GAA AAG PAG AAT TCC TCC TCC GGA ATC AGT GGG TCC			1703
Leu His Arg Gly Glu Lys Lys Asn Ser Ser Ser Gly Ile Ser Gly Ser			
355	360	365	
CAG CGG CGC TAC ATT GGG GTG ATG ATG CTG ACC CTG AGT CCC AGG GCT			1751
Gln Arg Arg Tyr Ile Gly Val Met Met Leu Thr Leu Ser Pro Arg Ala			
370	375	380	
GGT CTG CGG CCT GGT GAT GTG ATT TTG GGC ATT GGG GAG CAG ATG GTA			1799
Gly Leu Arg Pro Gly Asp Val Ile Leu Ala Ile Gly Glu Gln Met Val			
385	390	395	
CAA AAT GCT GAA GAT GTT TAT GAA GCT GTT CGA ACC CAA TCC CAG TTG			1847
Gln Asn Ala Glu Asp Val Tyr Glu Ala Val Arg Thr Gln Ser Gln Leu			
400	405	410	415
GCA GTG CAG ATC CGG CGG GGA CGA GAA ACA CTG ACC TTA TAT GTG ACC			1895
Ala Val Gln Ile Arg Arg Gly Arg Glu Thr Leu Thr Leu Tyr Val Thr			
420	425	430	
CCT GAG GTC ACA GAA TGAATAGATC ACCARGAGTA TGAGGCCTCT GCTCTGATTT CC			1952
Pro Glu Val Thr Glu			
435			

整理番号 156220

TCCTTGCCCTT TCTGGCTGAG GTTCTGAGGG CACCGAGACA GAGGGTTAAA TGAACCAGTG	2012
GGGGCAGGTC CCTCCAACCA CCAGCACTGA CTCCTGGGCT CTGAAGAATC ACAGAAACAC	2072
TTTTTATATA AATAAAPT ATACCTAGCA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA	2132
AAAAAAAAAA AA	2144

SEQ ID NO:29:

Met	Ala	Ala	Pro	Arg	Ala	Gly	Arg	Gly	Ala	Gly	Trp	Ser	Leu	Arg	Ala
1					5				10					15	
Trp	Arg	Ala	Leu	Gly	Gly	Ile	Arg	Trp	Gly	Arg	Arg	Pro	Arg	Leu	Thr
			20					25						30	
Pro	Asp	Leu	Arg	Ala	Leu	Leu	Thr	Ser	Gly	Thr	Ser	Asp	Pro	Arg	Ala
			35				40						45		
Arg	Val	Thr	Tyr	Gly	Thr	Pro	Ser	Leu	Trp	Ala	Arg	Leu	Ser	Val	Gly
			50				55					60			
Val	Thr	Glu	Pro	Arg	Ala	Cys	Leu	Thr	Ser	Gly	Thr	Pro	Gly	Pro	Arg
65					70					75				80	
Ala	Gln	Leu	Thr	Ala	Val	Thr	Pro	Asp	Thr	Arg	Thr	Arg	Glu	Ala	Ser
				85					90					95	
Glu	Asn	Ser	Gly	Thr	Arg	Ser	Arg	Ala	Trp	Leu	Ala	Val	Ala	Leu	Gly
			100					105						110	
Ala	Gly	Gly	Ala	Val	Leu	Leu	Leu	Leu	Trp	Gly	Gly	Gly	Arg	Gly	Pro
			115				120						125		
Pro	Ala	Val	Leu	Ala	Ala	Val	Pro	Ser	Pro	Pro	Pro	Ala	Ser	Pro	Arg
			130				135						140		
Ser	Gln	Tyr	Asn	Phe	Ile	Ala	Asp	Val	Val	Glu	Lys	Thr	Ala	Pro	Ala
145				150						155				160	
Val	Val	Tyr	Ile	Glu	Ile	Leu	Asp	Arg	His	Pro	Phe	Leu	Gly	Arg	Glu
				165					170					175	
Val	Pro	Ile	Ser	Asn	Gly	Ser	Gly	Phe	Val	Val	Ala	Ala	Asp	Gly	Leu
			180					185						190	
Ile	Val	Thr	Asn	Ala	His	Val	Val	Ala	Asp	Arg	Arg	Arg	Val	Arg	Val
			195					200					205		
Arg	Leu	Leu	Ser	Gly	Asp	Thr	Tyr	Glu	Ala	Val	Val	Thr	Ala	Val	Asp
			210				215					220			
Pro	Val	Ala	Asp	Ile	Ala	Thr	Leu	Arg	Ile	Gln	Thr	Lys	Glu	Pro	Leu
225				230						235				240	
Pro	Thr	Leu	Pro	Leu	Gly	Arg	Ser	Ala	Asp	Val	Arg	Gln	Gly	Glu	Phe
				245					250					255	
Val	Val	Ala	Met	Gly	Ser	Pro	Phe	Ala	Leu	Gln	Asn	Thr	Ile	Thr	Ser
			260					265					270		
Gly	Ile	Val	Ser	Ser	Ala	Gln	Arg	Pro	Ala	Arg	Asp	Leu	Gly	Leu	Pro
			275				280						285		

整理番号 156220

Gln Thr Asn Val Glu Tyr Ile Gln Thr Asp Ala Ala Ile Asp Phe Gly  
 290 295 300  
 Asn Ser Gly Gly Pro Leu Val Asn Leu Ala Arg Glu Leu Gly Ala Val  
 305 310 315 320  
 Ser Leu Gln Asp Gly Glu Val Ile Gly Val Asn Thr Met Lys Val Thr  
 325 330 335  
 Ala Gly Ile Ser Phe Ala Ile Pro Ser Asp Arg Leu Arg Glu Phe Leu  
 340 345 350  
 His Arg Gly Glu Lys Lys Asn Ser Ser Ser Gly Ile Ser Gly Ser Gln  
 355 360 365  
 Arg Arg Tyr Ile Gly Val Met Met Leu Thr Leu Ser Pro Arg Ala Gly  
 370 375 380  
 Leu Arg Pro Gly Asp Val Ile Leu Ala Ile Gly Glu Gln Met Val Gln  
 385 390 395 400  
 Asn Ala Glu Asp Val Tyr Glu Ala Val Arg Thr Gln Ser Gln Leu Ala  
 405 410 415  
 Val Gln Ile Arg Arg Gly Arg Glu Thr Leu Thr Leu Tyr Val Thr Pro  
 420 425 430  
 Glu-Val Thr Glu  
 435

SEQ ID NO:30:

FEATURE: Polymorphic variants at 672 and 1435

aa24=Arg/Cys aa278=Ala/Val

CGTGGATCCC GAGAAAGAGG CGCAGGACGA GGAGGCAGAA CCGACTGGC GCGTAGAGCA 60  
 GCAGCACGAG CAGTAGGAAG CAGTCACCCG GAAGCCCTGGG GGCGAGAGGC GAAGTGGTCA 120  
 GGCGCCGAAG GCCGAGAGCA CGCGGGGATC GGTCTCTTCC CGCCGGGTCT CTTACCGGTG 180  
 CGAGTCAAAG AGCCGCTCCG GCGCCGGCCC TGAGGGGARG TCATAACTG CTGCTTCAGG 240  
 AGCGCCCGGC CGTCGCGCGC GCGCCATT TCGCGCCCGG CGCAGGGGC TCTGGGAAG 300  
 GCGGAGTCTT TGGGCATCCG CCGGGGTGA GGGGACCCGA AGTCCTGAGG CGCGCCGGAA 360  
 GGGCTAGCGG TCOCAGCATA CCGCGCGGCC CTTGGGCGG TCTCACAACT CGCGTCCGGC 420  
 GGAGACCACA ATTCCCGCCA TTCGTGGGCG AGGGAGGAGT CCGCCTCCCG GAATCCTGGT 480  
 CCGGGCGTGC ACTTCTGAAG GACTTCAGGT ACCGGCGTGC CCGCGTCTCT ACTGTCCGCC 540  
 TGCTCGCGTC CTGGGTGCGG CCTCTGAGTA GGGCGGGCGA GGAGGCAGCC AAGGCGGAGC 600  
 TG ARG GCT GCG CCG AGG GCG GGG CGG GGT GCA GGC TGG AGC CTT CGG 647  
 Met Ala Ala Pro Arg Ala Gly Arg Gly Ala Gly Trp Ser Leu Arg  
 1 5 10 15  
 GCA TGG CGG GCT TTG GGG GGC ATT YGC TGG GGG AGG AGA CCC CGT TTG 695  
 Ala Trp Arg Ala Leu Gly Gly Ile Xaa Trp Gly Arg Arg Pro Arg Leu  
 20 25 30

整理番号 156220

ACC CCT GAC CTC CGG GCC CTG CTG ACG TCA GGA ACT TCT GAC CCC CGG	743
Thr Pro Asp Leu Arg Ala Leu Leu Thr Ser Gly Thr Ser Asp Pro Arg	
35 40 45	
GCC CGA GTG ACT TAT GGG ACC CCC AGT CTC TGG GCC CGG TTG TCT GTT	791
Ala Arg Val Thr Tyr Gly Thr Pro Ser Leu Trp Ala Arg Leu Ser Val	
50 55 60	
GGG GTC ACT GAA CCC CGA GCA TGC CTG ACG TCT GGG ACC CCG GGT CCC	839
Gly Val Thr Glu Pro Arg Ala Cys Leu Thr Ser Gly Thr Pro Gly Pro	
65 70 75	
CGG GCA CAA CTG ACT GCG GTG ACC CCA GAT ACC AGG ACC CGG GAG GCC	887
Arg Ala Gln Leu Thr Ala Val Thr Pro Asp Thr Arg Thr Arg Glu Ala	
80 85 90 95	
TCA GAG AAC TCT GGA ACC CGT TCG CGC GCG TGG CTG GCG GTG GCG CTG	935
Ser Glu Asn Ser Gly Thr Arg Ser Arg Ala Trp Leu Ala Val Ala Leu	
100 105 110	
GGC GCT GGG GGG GCA GTG CTG TTG TTG TTG TGG GGC GGG GGT CGG GGT	983
Gly Ala Gly Gly Ala Val Leu Leu Leu Leu Trp Gly Gly Gly Arg Gly	
115 120 125	
CCT CCG GCC GTC CTC GCC GCC GTC CCT AGC CCG CCG CCC GCT TCT CCC	1031
Pro Pro Ala Val Leu Ala Ala Val Pro Ser Pro Pro Pro Ala Ser Pro	
130 135 140	
CGG AGT CAG TAC AAC TTC ATC GCA GAT GTG GTG GAG AAG ACA GCA CCT	1079
Arg Ser Gln Tyr Asn Phe Ile Ala Asp Val Val Glu Lys Thr Ala Pro	
145 150 155	
GCC GTG GTC TAT ATC GAG ATC CTG GAC CGG CAC CCT TTC TTG GGC CGC	1127
Ala Val Val Tyr Ile Glu Ile Leu Asp Arg His Pro Phe Leu Gly Arg	
160 165 170 175	
GAG GTC CCT ATC TCG AAC GGC TCA GGA TTC GTG GTG GCT GCC GAT GGG	1175
Glu Val Pro Ile Ser Asn Gly Ser Gly Phe Val Val Ala Ala Asp Gly	
180 185 190	
CTC ATT GTC ACC AAC GCC CAT GTG GTG GCT GAT CGG CGC AGA GTC CGT	1223
Leu Ile Val Thr Asn Ala His Val Val Ala Asp Arg Arg Arg Val Arg	

整理番号 156220

195	200	205	
GTG AGA CTG CTA AGC GGC GAC ACG TAT GAG GCC GTG GTC ACA GCT GTG Val Arg Leu Leu Ser Gly Asp Thr Tyr Glu Ala Val Val Thr Ala Val 210 215 220			1271
GAT CCC GTG GCA GAC ATC GCA ACG CTG AGG ATT CAG ACT AAG GAG CCT Asp Pro Val Ala Asp Ile Ala Thr Leu Arg Ile Gln Thr Lys Glu Pro 225 230 235			1319
CTC CCC ACG CTG CCT CTG GGA CGC TCA GCT GAT GTC CGG CAA GGG GAG Leu Pro Thr Leu Pro Leu Gly Arg Ser Ala Asp Val Arg Gln Gly Glu 240 245 250 255			1367
TTT GTT GTT GCC ATG GGA AGT CCC TTT GCA CTG CAG AAC ACG ATC ACA Phe Val Val Ala Met Gly Ser Pro Phe Ala Leu Gln Asn Thr Ile Thr 260 265 270			1415
TCC GGC ATT GTT AGC TCT YCT CAG CGT CCA GCC AGA GAC CTG GGA CTC Ser Gly Ile Val Ser Ser Xaa Gln Arg Pro Ala Arg Asp Leu Gly Leu 275 280 285			1463
CCC CAA ACC AAT GTG GAA TAC ATT CAA ACT GAT GCA GCT ATT GAT TTT Pro Gln Thr Asn Val Glu Tyr Ile Gln Thr Asp Ala Ala Ile Asp Phe 290 295 300			1511
GGA AAC TCT GGA GGT CCC CTG GTT AAC CTG GAT GGG GAG GTG ATT GGA Gly Asn Ser Gly Gly Pro Leu Val Asn Leu Asp Gly Glu Val Ile Gly 305 310 315			1559
GTG AAC ACC ATG AAG GTC ACA GCT GGA ATC TCC TTT GCC ATC CCT TCT Val Asn Thr Met Lys Val Thr Ala Gly Ile Ser Phe Ala Ile Pro Ser 320 325 330 335			1607
GAT CGT CTT CGA GAG TTT CTG CAT CGT GGG GAA AAG AAG AAT TCC TCC Asp Arg Leu Arg Glu Phe Leu His Arg Gly Glu Lys Lys Asn Ser Ser 340 345 350			1655
TCC GGA ATC AGT GGG TCC CAG CGG CGC TAC ATT GGG GTG ATG ATG CTG Ser Gly Ile Ser Gly Ser Gln Arg Arg Tyr Ile Gly Val Met Met Leu 355 360 365			1703

整理番号 156220

ACC CTG AGT CCC AGC ATC CTT GCT GAA CTA CAG CTT CGA GAA CCA AGC 1751  
 Thr Leu Ser Pro Ser Ile Leu Ala Glu Leu Gln Leu Arg Glu Pro Ser  
 370 375 380

TTT CCC GAT GTT CAG CAT GGT GTA CTC ATC CAT AAA GTC ATC CTG GGC 1799  
 Phe Pro Asp Val Gln His Gly Val Leu Ile His Lys Val Ile Leu Gly  
 385 390 395

TCC CCT GCA CAC CGG GCT GGT CTC CGG CCT GGT GAT GTG ATT TTG GCC 1847  
 Ser Pro Ala His Arg Ala Gly Leu Arg Pro Gly Asp Val Ile Leu Ala  
 400 405 410 415

ATT GGG GAG CAG ATG GTA CAA AAT GCT GAA GAT GTT TAT GAA GCT GTT 1895  
 Ile Gly Glu Gln Met Val Gln Asn Ala Glu Asp Val Tyr Glu Ala Val  
 420 425 430

CGA ACC CAA TCC CAG TTG GCA GTG CAG ATC CGG CGG GGA CGA GAA ACA 1943  
 Arg Thr Gln Ser Gln Leu Ala Val Gln Ile Arg Arg Gly Arg Glu Thr  
 435 440 445

CTG ACC TTA TAT GTG ACC CCT GAG GTC ACA GAA TGAATAGATC ACCAAGAGTA 1996  
 Leu Thr Leu Tyr Val Thr Pro Glu Val Thr Glu  
 450 455

TGAGGCTCCT GCTCTGATT CCTCCTTGCC TTTCTGGCTG AGGTICTGAG GGCACCGAGA 2056  
 CAGAGGGTTA AATGAACCAG TGGGGGCAGG TCCCTCCAAC CACCAGCACT GACTCCTGGG 2116  
 CTCTGAAGAA TCACAGAAAC ACTTTTATA TAAATATAAA TTATACCTAG CAACATAAAA 2176  
 AAAAAAAAAA A 2187

SEQ ID NO:31:

Feature

24 Xaa = Arg or Cys

278 Xaa = Ala or Val

Met Ala Ala Pro Arg Ala Gly Arg Gly Ala Gly Trp Ser Leu Arg Ala  
 1 5 10 15  
 Trp Arg Ala Leu Gly Gly Ile Xaa Trp Gly Arg Arg Pro Arg Leu Thr  
 20 25 30  
 Pro Asp Leu Arg Ala Leu Leu Thr Ser Gly Thr Ser Asp Pro Arg Ala  
 35 40 45  
 Arg Val Thr Tyr Gly Thr Pro Ser Leu Trp Ala Arg Leu Ser Val Gly  
 50 55 60

逆組番号 156220

Val Thr Glu Pro Arg Ala Cys Leu Thr Ser Gly Thr Pro Gly Pro Arg			
65	70	75	80
Ala Gln Leu Thr Ala Val Thr Pro Asp Thr Arg Thr Arg Glu Ala Ser			
85	90		95
Glu Asn Ser Gly Thr Arg Ser Arg Ala Trp Leu Ala Val Ala Leu Gly			
100	105		110
Ala Gly Gly Ala Val Leu Leu Leu Leu Trp Gly Gly Gly Arg Gly Pro			
115	120		125
Pro Ala Val Leu Ala Ala Val Pro Ser Pro Pro Pro Ala Ser Pro Arg			
130	135		140
Ser Gln Tyr Asn Phe Ile Ala Asp Val Val Glu Lys Thr Ala Pro Ala			
145	150		155
Val Val Tyr Ile Glu Ile Leu Asp Arg His Pro Phe Leu Gly Arg Glu			
165	170		175
Val Pro Ile Ser Asn Gly Ser Gly Phe Val Val Ala Ala Asp Gly Leu			
180	185		190
Ile Val Thr Asn Ala His Val Val Ala Asp Arg Arg Arg Val Arg Val			
195	200		205
Arg Leu Leu Ser Gly Asp Thr Tyr Glu Ala Val Val Thr Ala Val Asp			
210	215		220
Pro Val Ala Asp Ile Ala Thr Leu Arg Ile Gln Thr Lys Glu Pro Leu			
225	230		235
Pro Thr Leu Pro Leu Gly Arg Ser Ala Asp Val Arg Gln Gly Glu Phe			
245	250		255
Val Val Ala Met Gly Ser Pro Phe Ala Leu Gln Asn Thr Ile Thr Ser			
260	265		270
Gly Ile Val Ser Ser Xaa Gln Arg Pro Ala Arg Asp Leu Gly Leu Pro			
275	280		285
Gln Thr Asn Val Glu Tyr Ile Gln Thr Asp Ala Ala Ile Asp Phe Gly			
290	295		300
Asn Ser Gly Gly Pro Leu Val Asn Leu Asp Gly Glu Val Ile Gly Val			
305	310		315
Asn Thr Met Lys Val Thr Ala Gly Ile Ser Phe Ala Ile Pro Ser Asp			
325	330		335
Arg Leu Arg Glu Phe Leu His Arg Gly Glu Lys Lys Asn Ser Ser Ser			
340	345		350
Gly Ile Ser Gly Ser Gln Arg Arg Tyr Ile Gly Val Met Met Leu Thr			
355	360		365
Leu Ser Pro Ser Ile Leu Ala Glu Leu Gln Leu Arg Glu Pro Ser Phe			
370	375		380
Pro Asp Val Gln His Gly Val Leu Ile His Lys Val Ile Leu Gly Ser			
385	390		400

登録番号 156220

Pro Ala His Arg Ala Gly Leu Arg Pro Gly Asp Val Ile Leu Ala Ile  
                     405                    410                    415  
 Gly Glu Gln Met Val Gln Asn Ala Glu Asp Val Tyr Glu Ala Val Arg  
                     420                    425                    430  
 Thr Gln Ser Gln Leu Ala Val Gln Ile Arg Arg Gly Arg Glu Thr Leu  
                     435                    440                    445  
 Thr Leu Tyr Val Thr Pro Glu Val Thr Glu  
                     450                    455

SEQ ID NO:32:

CATCCGGCAT TGTTAGCTCT GC 22

SEQ ID NO:33:

CATCCGGCAT TGTTAGCTCT GT 22

SEQ ID NO:34:

CAATAGCTGC ATCAGTTTGA ATG 23

SEQ ID NO:35:

TGGCGGGCTT TGGGGGGCAT TC 22

SEQ ID NO:36:

TGGCGGGCTT TGGGGGGCAT TT 22

SEQ ID NO:37:

GACGTCAGCA GGGCCCGGAG GTC 23

SEQ ID NO:38:

GATACCCAG CAGAAGCTGG 20

SEQ ID NO:39:

GATACCCAG CAGAAGCTGT 20

SEQ ID NO:40:

GCTGACATCA TTGGCGGAGA C 21



発明番号 156220

#### Claims

1. An isolated polynucleotide encoding a biologically active PSP1 polypeptide.
2. An isolated polynucleotide selected from the group consisting of:
  - (a) a polynucleotide encoding PSP1-1 having the nucleotide sequence as set forth in SEQ ID NO: 24 from nucleotide 603 to 1979;
  - (b) a polynucleotide encoding PSP1-2 having the nucleotide sequence as set forth in SEQ ID NO: 23 from nucleotide 603 to 1979;
  - (c) a polynucleotide encoding PSP1-3 having the nucleotide sequence as set forth in SEQ ID NO: 26 from nucleotide 603 to 1736;
  - (d) a polynucleotide encoding PSP1-4 having the nucleotide sequence as set forth in SEQ ID NO: 28 from nucleotide 603 to 1913; and
  - (e) a polynucleotide encoding D87257 (1325T) protein.
3. An isolated polynucleotide substantially similar to SEQ ID NO: 24; SEQ ID NO: 23; SEQ ID NO: 26; or SEQ ID NO: 28.
4. An isolated polynucleotide as claimed in claim 2 or 3 wherein nucleotides 672 and 1435 are independently selected from C and T.
5. An isolated polynucleotide having the nucleotide sequence as set forth in SEQ ID NOs: 23, 24, 26, 28, or 30 or SEQ ID NO: 17 wherein nucleotide 1325 is T.
6. A functional polypeptide encoded by the polynucleotide of any one of claims 1 to 5.
7. The functional polypeptide of claim 6 which is:  
PSP1-1 having the amino acid sequence set forth in SEQ ID NO: 25 or 30;  
PSP1-2 having the amino acid sequence set forth in SEQ ID NO: 8;  
PSP1-3 having the amino acid sequence set forth in SEQ ID NO: 27;  
PSP1-4 having the amino acid sequence set forth in SEQ ID NO: 29; or  
D87257 (1325T) protein having the amino acid sequence set forth in SEQ ID NO: 18  
wherein amino acid residue 213 is val.
8. The polynucleotide as claimed in of any one of claims 1 to 5 which is DNA or RNA.
9. A vector comprising the DNA of claim 8.

登録番号 1 5 6 2 2 0

10. A recombinant host cell comprising the vector of claim 9.
11. A method for preparing essentially pure PSP1 protein or D87257 (1325T) protein comprising culturing the recombinant host cell of claim 10 under conditions promoting expression of the protein and recovering the expressed protein.
12. PSP1 or D87257 (1325T) produced by the process of claim 11.
13. An antisense oligonucleotide comprising a sequence which is capable of binding to the polynucleotide of any one of claims 1 to 5 or D87258.
14. A modulator of the polypeptide of claim 6 or of D87258.
15. A method for assaying a medium for the presence of a substance that modulates PSP1 or D87258 activity:
  - (i) by affecting the binding of PSP1 or D87258 to cellular binding partners comprising the steps of:
    - (a) providing a PSP1 polypeptide of claim 8 or D87258 protein, or a functional derivative thereof and a cellular binding partner or synthetic analog thereof;
    - (b) incubating with a test substance which is suspected of modulating PSP1 or D87258 activity under conditions which permit the formation of a PSP1 or D87258 protein/cellular binding partner complex;
    - (c) assaying for the presence of the complex, free PSP1 or D87258 protein or free cellular binding partner; and
    - (d) comparing to a control to determine the effect of the substance;
  - (ii) by inhibiting proteolytic activity on a cellular substrate comprising the steps of:
    - (a) providing a PSP1 polypeptide of claim 8 or D87258, or a functional derivative thereof and a cellular substrate or synthetic analog thereof;
    - (b) incubating with a test substance which is suspected of inhibiting PSP1 or D87258 activity under conditions which permit the formation of a PSP1 or D87258 enzyme/substrate complex and subsequent cleavage of the substrate;
    - (c) assaying for the presence of proteolytically cleaved substrate; and
    - (d) comparing to a control to determine the effect of the substance.
  - (iii) by direct binding to PSP1 or D87258 protein comprising the steps of:

整理番号 156220

- (a) providing a labelled PSP1 polypeptide of claim 8 or D87258, or a functional derivative thereof;
  - (b) providing solid support-associated modulator candidates;
  - (c) incubating a mixture of the labelled PSP1 or D87258 protein with the support-associated modulator candidates under conditions which can permit the formation of a PSP1 or D87258 protein/modulator candidate complex;
  - (d) separating the solid support from free soluble-labelled PSP1 or D87258 protein;
  - (e) assaying for the presence of solid support-associated labelled protein;
  - (f) isolating the solid support complexed with labelled PSP1 or D87258 protein;
- and
- (g) identifying the modulator candidate.

16. PSP1 or D87258 protein modulating compounds identified by the method of claim 15.

17. The use of a modulating compound of claim 16 in the manufacture of a medicament for treating of a patient having need to modulate PSP1 or D87258 activity.

18. A pharmaceutical composition comprising the modulating compound of claim 16 and a pharmaceutically acceptable carrier.

19. A method of diagnosing conditions associated with PSP1 or D87258 protein deficiency which comprises:

- (a) isolating a polynucleotide sample from an individual;
- (b) assaying the polynucleotide sample and a polynucleotide encoding PSP1 having the nucleotide sequence as set forth in SEQ ID NOs: 23, 24, 26, 28 or 30 or a polynucleotide encoding D87258 as set forth in SEQ ID NO: 18; and
- (c) comparing differences between the polynucleotide sample and the PSP1 or D87258 polynucleotide, wherein any differences indicate mutations in the PSP1 or D87258 sequence.

20. A method of treating conditions which are related to insufficient PSP1 or D87258 protein function which comprises:

- (i) the steps of:
  - (a) isolating cells from a patient deficient in PSP1 or D87258 protein function;
  - (b) altering the cells by transfecting the polynucleotide of any one of claims 1 to 7, or a polynucleotide encoding D87258 as set forth in SEQ ID NO: 18 into the cells wherein a PSP1 or D87258 protein is expressed; and

整理番号 156220

(c) introducing the cells back to the patient to alleviate the condition; or  
(ii) administering the polynucleotide of any one of claims 1 to 5, or a polynucleotide encoding D87258 as set forth in SEQ ID NO: 18, to a patient deficient in PSP1 or D87258 protein function wherein a PSP1 or D87258 protein is expressed and alleviates the condition.

21. An antibody immunoreactive with PSP1, D87258 or an immunogen thereof.

22. A transgenic non-human animal capable of expressing in any cell thereof the DNA of claim 5 or a polynucleotide encoding D87258 as set forth in SEQ ID NO: 18.

23. A method for determining the genetic predisposition to neurodegeneration in a patient comprising detecting PSP1 or D87258 polymorphisms in a sample from a patient, preferably neurodegeneration predisposition to Alzheimer's disease.

24. The method of claim 23 wherein the polymorphisms detected are at nucleotide 672 of PSP1, at nucleotide 1435 of PSP1 or at nucleotide 1325 of D87258.

25. The method of claim 24 wherein the polymorphisms are detected by polymerase chain reaction, preferably wherein the oligonucleotides used with the polymerase chain reaction have a nucleotide sequence selected from the group consisting of SEQ ID NOs: 32, 33, 34, 35, 36, 37, 38, 39, or 40.

26. An isolated polynucleotide having the nucleotide sequence as set forth in SEQ ID NO: 32, 33, 34, 35, 36, 37, 38, 39, or 40.

27. An oligonucleotide pair comprising oligonucleotides having the nucleotide sequence as set forth in:

- (a) SEQ ID NOs: 32 and 34;
- (b) SEQ ID NOs: 33 and 34;
- (c) SEQ ID NOs: 35 and 37;
- (d) SEQ ID NOs: 36 and 37;
- (e) SEQ ID NOs: 38 and 40; or
- (f) SEQ ID NOs: 39 and 40.

41 SCTSDPRARVTYGTPLWARLSVGVTEPRACLTSGTGPRAQLTAVTFDT 90  
 . | . : . | : : . : . | : : . | : . | . . | . : . : .  
 2 KKTTLALSRLALSGLALSPLSATAAETSSATTAQQMPSLAPMLEKVMP 51  
 91 RTREASENSGTRSRRAWLAVLGGAGVLLLLWGGGRGPPAVLAAPVSPFP 140  
 . . . | . | . : : : . : . : : : . : : . : . . | .  
 52 VVSINVEGSTTVNTPRMPRNEQQ . . . . . FFGDD . . . . SPFCQEGSPFQ 90  
 141 ASPRSQYNFIADVVEKTAPAVVYIEILDRHPFLGREVPISNGSGFVVAAD 190  
 . | | : | : : : . : . : : : . : : . | | : : . | |  
 91 SSPFCQ . . . GGQGGNGGGQQQKFHAL . . . . . GSGVIIDAD 122  
 191 . GLIVTNAHVVDARRVRVRLSGDYEAVVTAVDPVADIATLRIQTKEP 239  
 | . : | | . | | : : | . | : : | : : | | . | | : | | . .  
 123 KGYVVTNNHVDNATVIKVLQSOGRKFDAKMGKOPRSOIALIQIQNPKN 172  
 240 LPTLPLGRSADVRQGEFVVAMGSPFALQNTITSGIVSSAQRPARDLGLPQ 289  
 | . : : : | . : | | : : | : | : | : | : | : | . | . | | .  
 173 LTAIKMADSALRVGDYTVGIGNPFGLGETVTSGIVSALGRS . . . GLNA 218  
 290 TNVE . YIQTDAIDFCNSGGPLVNLGEGVIGVNTMKVTA . . . . GISFAI 333  
 . | | : | | | | : | | | | : | | | : | | | : : : | | : | |  
 219 ENYENFIQTDAAINRCNSGGALVNLNGELIGINTAILAPDGGNIGIGFAI 268  
 334 PSDRLREF . . . . . LHRGE . . . . . 346  
 | | : : : : : : : | |  
 269 PSNMVKNLTSQMVEYGVQVKGELGIMGTELNSELAKAMKVDAQRGA FVSQ 318  
 347 . KKNSSSGISG . . . . . SQRRYIGVM . . . . MLTL . . . 369  
 . | | | : : | . : | : | . | | . | |  
 319 VLPNSSAAKAGIKAGDVITSLNGKPISSFAALRAQVGTMVPVGSKLT LGLL 368  
 370 . . . . . SPSILAEQLREPSFPDVQHGVLIHKVIL 398  
 | . | | : : : : | | : : : | : : : |  
 369 RDGKQVNVNLELQQSSQNQVDSSSIFNGIEGAEMSNKGKDQGVVNVNVKT 418  
 399 GSPAHRAGLRPGOVILAIGEOMVQNAEDVYEAVRTQ . SOLAVQIRRGRET 447  
 | . | | . | | : | | | : : : | | . | : : . . | | : | |  
 419 GTPAAQIGLKKGDVIIGANQQAVKNIAELRKVLDSKPSVLALNIQRGDRH 468  
 448 LTLVTFEVE 458  
 | . : . : . :  
 469 LPVNAVISLNP 479

FIGURE 1/3

整理番号 156220

ページ(2)

	1		50
PSP1-2	CGTGGATCCC GAGAAAGAGG CGCAGGACGA GGAGGCAGAA CCCGACTGGC		
PSP1-1	CGTGGATCCC GAGAAAGAGG CGCAGGACGA GGAGGCAGAA CCCGACTGGC		
PSP1-3	CGTGGATCCC GAGAAAGAGG CGCAGGACGA GGAGGCAGAA CCCGACTGGC		
PSP1-4	CGTGGATCCC GAGAAAGAGG CGCAGGACGA GGAGGCAGAA CCCGACTGGC		
	51		100
PSP1-2	GCGTAGAGCA GCAGCACGAG CAGTAGGAAG CAGTCACCCG GAAGCCTGGG		
PSP1-1	GCGTAGAGCA GCAGCACGAG CAGTAGGAAG CAGTCACCCG GAAGCCTGGG		
PSP1-3	GCGTAGAGCA GCAGCACGAG CAGTAGGAAG CAGTCACCCG GAAGCCTGGG		
PSP1-4	GCGTAGAGCA GCAGCACGAG CAGTAGGAAG CAGTCACCCG GAAGCCTGGG		
	101		150
PSP1-2	GGCGAGAGGC GAAGTGGTCA GGCGCCGAAG GCCGAGAGCA CGCGGGGATC		
PSP1-1	GGCGAGAGGC GAAGTGGTCA GGCGCCGAAG GCCGAGAGCA CGCGGGGATC		
PSP1-3	GGCGAGAGGC GAAGTGGTCA GGCGCCGAAG GCCGAGAGCA CGCGGGGATC		
PSP1-4	GGCGAGAGGC GAAGTGGTCA GGCGCCGAAG GCCGAGAGCA CGCGGGGATC		
	151		200
PSP1-2	GGTCTCTTCC CGCCGGGTCT CTTACCGGTG CGAGTCAAAG AGCCGCTCCG		
PSP1-1	GGTCTCTTCC CGCCGGGTCT CTTACCGGTG CGAGTCAAAG AGCCGCTCCG		
PSP1-3	GGTCTCTTCC CGCCGGGTCT CTTACCGGTG CGAGTCAAAG AGCCGCTCCG		
PSP1-4	GGTCTCTTCC CGCCGGGTCT CTTACCGGTG CGAGTCAAAG AGCCGCTCCG		
	201		250
PSP1-2	GCCCCGGCCC TGAGGGAAGC TCCATAACTG CTGCTTCAGG AGCGCCCGGC		
PSP1-1	GCCCCGGCCC TGAGGGAAGC TCCATAACTG CTGCTTCAGG AGCGCCCGGC		
PSP1-3	GCCCCGGCCC TGAGGGAAGC TCCATAACTG CTGCTTCAGG AGCGCCCGGC		
PSP1-4	GCCCCGGCCC TGAGGGAAGC TCCATAACTG CTGCTTCAGG AGCGCCCGGC		
	251		300
PSP1-2	CGTCGCCGCC GCCGCCATTT TCGCGCCCGG CCGCAGGGGC TCTTGGGAAG		
PSP1-1	CGTCGCCGCC GCCGCCATTT TCGCGCCCGG CCGCAGGGGC TCTTGGGAAG		
PSP1-3	CGTCGCCGCC GCCGCCATTT TCGCGCCCGG CCGCAGGGGC TCTTGGGAAG		
PSP1-4	CGTCGCCGCC GCCGCCATTT TCGCGCCCGG CCGCAGGGGC TCTTGGGAAG		

FIGURE 2A/3

発明番号 156220

ページ(3)

	301		350
PSP1-2	GCGGAGTCTT TGGGCATCCG CCCGGGGTGA GGGGACCCGA AGTCCTGAGG		
PSP1-1	GCGGAGTCTT TGGGCATCCG CCCGGGGTGA GGGGACCCGA AGTCCTGAGG		
PSP1-3	GCGGAGTCTT TGGGCATCCG CCCGGGGTGA GGGGACCCGA AGTCCTGAGG		
PSP1-4	GCGGAGTCTT TGGGCATCCG CCCGGGGTGA GGGGACCCGA AGTCCTGAGG		
	351		400
PSP1-2	CGCGCCGGAA GGGCTAGCGG TOCCAGCATA CCCC GCGGCC CCTTGGGCCG		
PSP1-1	CGCGCCGGAA GGGCTAGCGG TOCCAGCATA CCCC GCGGCC CCTTGGGCCG		
PSP1-3	CGCGCCGGAA GGGCTAGCGG TOCCAGCATA CCCC GCGGCC CCTTGGGCCG		
PSP1-4	CGCGCCGGAA GGGCTAGCGG TOCCAGCATA CCCC GCGGCC CCTTGGGCCG		
	401		450
PSP1-2	TCTCACAAC TCGCTCCGGC GGAGACCACA ATTCCCGGCA TTCGTGGGGC		
PSP1-1	TCTCACAAC TCGCTCCGGC GGAGACCACA ATTCCCGGCA TTCGTGGGGC		
PSP1-3	TCTCACAAC TCGCTCCGGC GGAGACCACA ATTCCCGGCA TTCGTGGGGC		
PSP1-4	TCTCACAAC TCGCTCCGGC GGAGACCACA ATTCCCGGCA TTCGTGGGGC		
	451		500
PSP1-2	AGGGAGGAGT CGGCCTCCCG GAATCCTGGT CCCGGCGTGC ACTTCTGAAG		
PSP1-1	AGGGAGGAGT CGGCCTCCCG GAATCCTGGT CCCGGCGTGC ACTTCTGAAG		
PSP1-3	AGGGAGGAGT CGGCCTCCCG GAATCCTGGT CCCGGCGTGC ACTTCTGAAG		
PSP1-4	AGGGAGGAGT CGGCCTCCCG GAATCCTGGT CCCGGCGTGC ACTTCTGAAG		
	501		550
PSP1-2	GACTTCAGGT ACCGGCGTGC CCCGCGTCCT ACTGTCCGCC TGCTCGCGTC		
PSP1-1	GACTTCAGGT ACCGGCGTGC CCCGCGTCCT ACTGTCCGCC TGCTCGCGTC		
PSP1-3	GACTTCAGGT ACCGGCGTGC CCCGCGTCCT ACTGTCCGCC TGCTCGCGTC		
PSP1-4	GACTTCAGGT ACCGGCGTGC CCCGCGTCCT ACTGTCCGCC TGCTCGCGTC		
	551		600
PSP1-2	CTGGGTGCCG CCTCTGAGTA GGGCGGGCGA GGAGGCAGCC AAGGCGGAGC		
PSP1-1	CTGGGTGCCG CCTCTGAGTA GGGCGGGCGA GGAGGCAGCC AAGGCGGAGC		
PSP1-3	CTGGGTGCCG CCTCTGAGTA GGGCGGGCGA GGAGGCAGCC AAGGCGGAGC		
PSP1-4	CTGGGTGCCG CCTCTGAGTA GGGCGGGCGA GGAGGCAGCC AAGGCGGAGC		

FIGURE 2B/3

整理番号 156220

ページ(4)

	601		650
PSP1-2	TGATGGCTGC	GGCGGGGGTG	CAGGCTGGAG CCTTCGGGCA
PSP1-1	TGATGGCTGC	GGCGGGGGTG	CAGGCTGGAG CCTTCGGGCA
PSP1-3	TGATGGCTGC	GGCGGGGGTG	CAGGCTGGAG CCTTCGGGCA
PSP1-4	TGATGGCTGC	GGCGGGGGTG	CAGGCTGGAG CCTTCGGGCA
	651		700
PSP1-2	TGGCGGGCTT	TGGGGGGCAT	TTGCTGGGGG AGGAGACCCC GTTTGACCCC
PSP1-1	TGGCGGGCTT	TGGGGGGCAT	TCGCTGGGGG AGGAGACCCC GTTTGACCCC
PSP1-3	TGGCGGGCTT	TGGGGGGCAT	TCGCTGGGGG AGGAGACCCC GTTTGACCCC
PSP1-4	TGGCGGGCTT	TGGGGGGCAT	TCGCTGGGGG AGGAGACCCC GTTTGACCCC
	701		750
PSP1-2	TGACCTCCGG	GCCCTGCTGA	CGTCAGGAAC TTCTGACCCC CGGGCCCGAG
PSP1-1	TGACCTCCGG	GCCCTGCTGA	CGTCAGGAAC TTCTGACCCC CGGGCCCGAG
PSP1-3	TGACCTCCGG	GCCCTGCTGA	CGTCAGGAAC TTCTGACCCC CGGGCCCGAG
PSP1-4	TGACCTCCGG	GCCCTGCTGA	CGTCAGGAAC TTCTGACCCC CGGGCCCGAG
	751		800
PSP1-2	TGACTTATGG	GACCCCCAGT	CTCTGGGCCC GGTGTCTGT TGGGCTCACT
PSP1-1	TGACTTATGG	GACCCCCAGT	CTCTGGGCCC GGTGTCTGT TGGGCTCACT
PSP1-3	TGACTTATGG	GACCCCCAGT	CTCTGGGCCC GGTGTCTGT TGGGCTCACT
PSP1-4	TGACTTATGG	GACCCCCAGT	CTCTGGGCCC GGTGTCTGT TGGGCTCACT
	801		850
PSP1-2	GAACCCCGAG	CATGCCTGAC	GTCTGGGACC CCGGGTCCCC GGGCACAAC
PSP1-1	GAACCCCGAG	CATGCCTGAC	GTCTGGGACC CCGGGTCCCC GGGCACAAC
PSP1-3	GAACCCCGAG	CATGCCTGAC	GTCTGGGACC CCGGGTCCCC GGGCACAAC
PSP1-4	GAACCCCGAG	CATGCCTGAC	GTCTGGGACC CCGGGTCCCC GGGCACAAC
	851		900
PSP1-2	GACTGCGGTG	AOCCAGATA	CCAGGACCCG GGAGGCTCA GAGAAGTCTG
PSP1-1	GACTGCGGTG	AOCCAGATA	CCAGGACCCG GGAGGCTCA GAGAAGTCTG
PSP1-3	GACTGCGGTG	AOCCAGATA	CCAGGACCCG GGAGGCTCA GAGAAGTCTG
PSP1-4	GACTGCGGTG	AOCCAGATA	CCAGGACCCG GGAGGCTCA GAGAAGTCTG

FIGURE 2C/3



整理番号 156220

ページ(5)

	901		950
PSP1-2	GAACCOGTTT	GCGCGCGTGG	CTGGCGGTGG CGCTGGGCGC TGGGGGGGCA
PSP1-1	GAACCOGTTT	GCGCGCGTGG	CTGGCGGTGG CGCTGGGCGC TGGGGGGGCA
PSP1-3	GAACCOGTTT	GCGCGCGTGG	CTGGCGGTGG CGCTGGGCGC TGGGGGGGCA
PSP1-4	GAACCOGTTT	GCGCGCGTGG	CTGGCGGTGG CGCTGGGCGC TGGGGGGGCA
	951		1000
PSP1-2	GTGCTGTTGT	TGTTGTGGGG	CGGGGGTCGG GGTCTCCGG CCGTCCTCGC
PSP1-1	GTGCTGTTGT	TGTTGTGGGG	CGGGGGTCGG GGTCTCCGG CCGTCCTCGC
PSP1-3	GTGCTGTTGT	TGTTGTGGGG	CGGGGGTCGG GGTCTCCGG CCGTCCTCGC
PSP1-4	GTGCTGTTGT	TGTTGTGGGG	CGGGGGTCGG GGTCTCCGG CCGTCCTCGC
	1001		1050
PSP1-2	CGCCGTCCCT	AGCCCGCCGC	CCGCTTCTCC CCGGAGTCAG TACAACTTCA
PSP1-1	CGCCGTCCCT	AGCCCGCCGC	CCGCTTCTCC CCGGAGTCAG TACAACTTCA
PSP1-3	CGCCGTCCCT	AGCCCGCCGC	CCGCTTCTCC CCGGAGTCAG TACAACTTCA
PSP1-4	CGCCGTCCCT	AGCCCGCCGC	CCGCTTCTCC CCGGAGTCAG TACAACTTCA
	1051		1100
PSP1-2	TCGCAGATGT	GGTGGAGAAG	ACAGCACCTG CCGTGGTCTA TATCGAGATC
PSP1-1	TCGCAGATGT	GGTGGAGAAG	ACAGCACCTG CCGTGGTCTA TATCGAGATC
PSP1-3	TCGCAGATGT	GGTGGAGAAG	ACAGCACCTG CCGTGGTCTA TATCGAGATC
PSP1-4	TCGCAGATGT	GGTGGAGAAG	ACAGCACCTG CCGTGGTCTA TATCGAGATC
	1101		1150
PSP1-2	CTGGACGGC	ACCCTTCTT	GGGCCGCGAG GTCCCTATCT CGAACGGCTC
PSP1-1	CTGGACGGC	ACCCTTCTT	GGGCCGCGAG GTCCCTATCT CGAACGGCTC
PSP1-3	CTGGACGGC	ACCCTTCTT	GGGCCGCGAG GTCCCTATCT CGAACGGCTC
PSP1-4	CTGGACGGC	ACCCTTCTT	GGGCCGCGAG GTCCCTATCT CGAACGGCTC
	1151		1200
PSP1-2	AGGATTCTGT	GTGGCTGCCG	ATGGGCTCAT TGTCACCAAC GCCCATGTGG
PSP1-1	AGGATTCTGT	GTGGCTGCCG	ATGGGCTCAT TGTCACCAAC GCCCATGTGG
PSP1-3	AGGATTCTGT	GTGGCTGCCG	ATGGGCTCAT TGTCACCAAC GCCCATGTGG
PSP1-4	AGGATTCTGT	GTGGCTGCCG	ATGGGCTCAT TGTCACCAAC GCCCATGTGG

FIGURE 2D/3

整理番号 156220

ページ(6)

	1201		1250
PSP1-2	TGGCTGATCG GCGCAGAGTC CGTGTGAGAC TGCTAAGCGG CGACACGTAT		
PSP1-1	TGGCTGATCG GCGCAGAGTC CGTGTGAGAC TGCTAAGCGG CGACACGTAT		
PSP1-3	TGGCTGATCG GCGCAGAGTC CGTGTGAGAC TGCTAAGCGG CGACACGTAT		
PSP1-4	TGGCTGATCG GCGCAGAGTC CGTGTGAGAC TGCTAAGCGG CGACACGTAT		
	1251		1300
PSP1-2	GAGGCCGTGG TCACAGCTGT GGATCCCGTG GCAGACATCG CAACGCTGAG		
PSP1-1	GAGGCCGTGG TCACAGCTGT GGATCCCGTG GCAGACATCG CAACGCTGAG		
PSP1-3	GAGGCCGTGG TCACAGCTGT GGATCCCGTG GCAGACATCG CAACGCTGAG		
PSP1-4	GAGGCCGTGG TCACAGCTGT GGATCCCGTG GCAGACATCG CAACGCTGAG		
	1301		1350
PSP1-2	GATTCAGACT AAGGAGCCTC TCCCCACGCT GCCTCTGGGA CGCTCAGCTG		
PSP1-1	GATTCAGACT AAGGAGCCTC TCCCCACGCT GCCTCTGGGA CGCTCAGCTG		
PSP1-3	GATTCAGACT AAGGAGCCTC TCCCCACGCT GCCTCTGGGA CGCTCAGCTG		
PSP1-4	GATTCAGACT AAGGAGCCTC TCCCCACGCT GCCTCTGGGA CGCTCAGCTG		
	1351		1400
PSP1-2	ATGTCCGGCA AGGGGAGTTT GTTGTGCGCA TGGGAAGTCC CTTTGCACTG		
PSP1-1	ATGTCCGGCA AGGGGAGTTT GTTGTGCGCA TGGGAAGTCC CTTTGCACTG		
PSP1-3	ATGTCCGGCA AGGGGAGTTT GTTGTGCGCA TGGGAAGTCC CTTTGCACTG		
PSP1-4	ATGTCCGGCA AGGGGAGTTT GTTGTGCGCA TGGGAAGTCC CTTTGCACTG		
	1401		1450
PSP1-2	CAGAACACGA TCACATCCGG CATTGTTAGC TCTGCTCAGC GTCCAGCCAG		
PSP1-1	CAGAACACGA TCACATCCGG CATTGTTAGC TCTGCTCAGC GTCCAGCCAG		
PSP1-3	CAGAACACGA TCACATCCGG CATTGTTAGC TCTGCTCAGC GTCCAGCCAG		
PSP1-4	CAGAACACGA TCACATCCGG CATTGTTAGC TCTGCTCAGC GTCCAGCCAG		
	1451		1500
PSP1-2	AGACCTGGGA CTCGCCAAA CCAATGTGGA ATACATTCAA ACTGATGCAG		
PSP1-1	AGACCTGGGA CTCGCCAAA CCAATGTGGA ATACATTCAA ACTGATGCAG		
PSP1-3	AGACCTGGGA CTCGCCAAA CCAATGTGGA ATACATTCAA ACTGATGCAG		
PSP1-4	AGACCTGGGA CTCGCCAAA CCAATGTGGA ATACATTCAA ACTGATGCAG		

FIGURE 2E/3

整理番号 156220

ページ(7)

	1501		1550
PSP1-2	CTATTGATTT TGGAAACTCT GGAGGTCCCC TGGTTAACCT .....		
PSP1-1	CTATTGATTT TGGAAACTCT GGAGGTCCCC TGGTTAACCT .....		
PSP1-3	CTATTGATTT TGGAAACTCT GGAGGTCCCC TGGTTAACCT GGTGAGTGAG		
PSP1-4	CTATTGATTT TGGAAACTCT GGAGGTCCCC TGGTTAACCT .....		
	1551		1600
PSP1-2	.....		
PSP1-1	.....		
PSP1-3	ACATCCTTCC TTCCAAGAAT CCCTGCCCCA GGTCAGTGIG GGAAGGGTAG		
PSP1-4	.....		
	1601		1650
PSP1-2	.....		
PSP1-1	.....		
PSP1-3	GTTTCCCCTA ATTCAAGGAT GTTGGTCAA GTTCTGAGC AGTTCTTTGT		
PSP1-4	.....		
	1651		1700
PSP1-2	.....		
PSP1-1	.....		
PSP1-3	TGGCTATCTC TCAATATCCA ACCAGATCTC CCCAACACTT GCTGGTACTT		
PSP1-4	.....		
	1701		1750
PSP1-2	.....		
PSP1-1	.....		
PSP1-3	TTGTTCGGGT GCCCCATCC CCTACTATTT GTT <u>TAGG</u> CTA GGGAACTGGG		
PSP1-4	.....GGCTA GGGAACTGGG		
	1751		1800
PSP1-2	.....GGATG GGGAGGTGAT TGGAGTGAAC ACCATGAAGG		
PSP1-1	.....GGATG GGGAGGTGAT TGGAGTGAAC ACCATGAAGG		
PSP1-3	GGCTGTATCC CTGCAGGATG GGGAGGTGAT TGGAGTGAAC ACCATGAAGG		
PSP1-4	GGCTGTATCC CTGCAGGATG GGGAGGTGAT TGGAGTGAAC ACCATGAAGG		

FIGURE 2F/3

整理番号 156220

ページ(8)

	1801		1850
PSP1-2	TCACAGCTGG AATCTCCTTT GCCATCCCTT CTGATCGTCT TCGAGAGTTT		
PSP1-1	TCACAGCTGG AATCTCCTTT GCCATCCCTT CTGATCGTCT TCGAGAGTTT		
PSP1-3	TCACAGCTGG AATCTCCTTT GCCATCCCTT CTGATCGTCT TCGAGAGTTT		
PSP1-4	TCACAGCTGG AATCTCCTTT GCCATCCCTT CTGATCGTCT TCGAGAGTTT		
	1851		1900
PSP1-2	CTGCATCGTG GGGAAAAGAA GAATTCCTCC TCCGGAATCA GTGGGTCCCA		
PSP1-1	CTGCATCGTG GGGAAAAGAA GAATTCCTCC TCCGGAATCA GTGGGTCCCA		
PSP1-3	CTGCATCGTG GGGAAAAGAA GAATTCCTCC TCCGGAATCA GTGGGTCCCA		
PSP1-4	CTGCATCGTG GGGAAAAGAA GAATTCCTCC TCCGGAATCA GTGGGTCCCA		
	1901		1950
PSP1-2	GCGGCGCTAC ATTGGGGTGA TGATGCTGAC CCTGAGTCCC AGCATCCTTG		
PSP1-1	GCGGCGCTAC ATTGGGGTGA TGATGCTGAC CCTGAGTCCC AGCATCCTTG		
PSP1-3	GCGGCGCTAC ATTGGGGTGA TGATGCTGAC CCTGAGTCCC AGCATCCTTG		
PSP1-4	GCGGCGCTAC ATTGGGGTGA TGATGCTGAC CCTGAGTCCC A.....		
	1951		2000
PSP1-2	CTGAACTACA GCTTCGAGAA CCAAGCTTTC CCGATGTTCA GCATGGTGTA		
PSP1-1	CTGAACTACA GCTTCGAGAA CCAAGCTTTC CCGATGTTCA GCATGGTGTA		
PSP1-3	CTGAACTACA GCTTCGAGAA CCAAGCTTTC CCGATGTTCA GCATGGTGTA		
PSP1-4	.....		
	2001		2050
PSP1-2	CTCATCCATA AAGTCATCCT GGGCTCCOCT GCACACCGGG CTGGTCTGCG		
PSP1-1	CTCATCCATA AAGTCATCCT GGGCTCCOCT GCACACCGGG CTGGTCTGCG		
PSP1-3	CTCATCCATA AAGTCATCCT GGGCTCCOCT GCACACCGGG CTGGTCTGCG		
PSP1-4	.....GGG CTGGTCTGCG		
	2051		2100
PSP1-2	GCCTGGTGAT GTGATTTTGG CCATTGGGGA GCAGATGGTA CAAAATGCTG		
PSP1-1	GCCTGGTGAT GTGATTTTGG CCATTGGGGA GCAGATGGTA CAAAATGCTG		
PSP1-3	GCCTGGTGAT GTGATTTTGG CCATTGGGGA GCAGATGGTA CAAAATGCTG		
PSP1-4	GCCTGGTGAT GTGATTTTGG CCATTGGGGA GCAGATGGTA CAAAATGCTG		

FIGURE 2G/3

整理番号 156220

ページ(9)

2101	2150
PSP1-2 AAGATGTTTA TGAAGCTGTT CGAACCCAAT CCCAGTTGGC AGTGCAGATC	
PSP1-1 AAGATGTTTA TGAAGCTGTT CGAACCCAAT CCCAGTTGGC AGTGCAGATC	
PSP1-3 AAGATGTTTA TGAAGCTGTT CGAACCCAAT CCCAGTTGGC AGTGCAGATC	
PSP1-4 AAGATGTTTA TGAAGCTGTT CGAACCCAAT CCCAGTTGGC AGTGCAGATC	
2151	2200
PSP1-2 CGGOGGGGAC GAGAAACACT GACCTTATAT GTGACCCCTG AGGTCACAGA	
PSP1-1 CGGOGGGGAC GAGAAACACT GACCTTATAT GTGACCCCTG AGGTCACAGA	
PSP1-3 CGGOGGGGAC GAGAAACACT GACCTTATAT GTGACCCCTG AGGTCACAGA	
PSP1-4 CGGOGGGGAC GAGAAACACT GACCTTATAT GTGACCCCTG AGGTCACAGA	
2201	2250
PSP1-2 ATGAATAGAT CACCAAGAGT ATGAGGCTCC TGCTCTGATT TCCTCCTTGC	
PSP1-1 ATGAATAGAT CACCAAGAGT ATGAGGCTCC TGCTCTGATT TCCTCCTTGC	
PSP1-3 ATGAATAGAT CACCAAGAGT ATGAGGCTCC TGCTCTGATT TCCTCCTTGC	
PSP1-4 ATGAATAGAT CACCAAGAGT ATGAGGCTCC TGCTCTGATT TCCTCCTTGC	
2251	2300
PSP1-2 CTTTCTGGCT GAGGTTCCTGA GGGCACCGAG ACAGAGGGTT AAATGAACCA	
PSP1-1 CTTTCTGGCT GAGGTTCCTGA GGGCACCGAG ACAGAGGGTT AAATGAACCA	
PSP1-3 CTTTCTGGCT GAGGTTCCTGA GGGCACCGAG ACAGAGGGTT AAATGAACCA	
PSP1-4 CTTTCTGGCT GAGGTTCCTGA GGGCACCGAG ACAGAGGGTT AAATGAACCA	
2301	2350
PSP1-2 GTGGGGGCAG GTCCCTCCAA CCACCAGCAC TGACTCCTGG GCTCTGAAGA	
PSP1-1 GTGGGGGCAG GTCCCTCCAA CCACCAGCAC TGACTCCTGG GCTCTGAAGA	
PSP1-3 GTGGGGGCAG GTCCCTCCAA CCACCAGCAC TGACTCCTGG GCTCTGAAGA	
PSP1-4 GTGGGGGCAG GTCCCTCCAA CCACCAGCAC TGACTCCTGG GCTCTGAAGA	
2351	2400
PSP1-2 ATCACAGAAA CACTTTTTAT ATAAAATAAA ATTATACCTA GCAACATAAA	
PSP1-1 ATCACAGAAA CACTTTTTAT ATAAAATAAA ATTATACCTA GCAACATAAA	
PSP1-3 ATCACAGAAA CACTTTTTAT ATAAAATAAA ATTATACCTA GCAACATATT	
PSP1-4 ATCACAGAAA CACTTTTTAT ATAAAATAAA ATTATACCTA GCAAAAAAAA	

FIGURE 2H/3

整理番号 156220

ページ(10)

	2401		2450
PSP1-2	AAAAAAAAAA AA.....		
PSP1-1	AAAAAAAAAA AA.....		
PSP1-3	ATAGTAAAA ATGAGGTGGG AGGGCTGGAT CTTTCCCCC ACCAAAAGGC		
PSP1-4	AAAAAAAAAA AAAAAAAAAA PAAAAAAPA AAAAA.....		
	2451		2500
PSP1-2	.....		
PSP1-1	.....		
PSP1-3	TAGAGGTAAA GCTGTATCCC CCTAACTTA GGGGAGATAC TGGAGCTGAC		
PSP1-4	.....		
	2501		2550
PSP1-2	.....		
PSP1-1	.....		
PSP1-3	CATCCTGACC TCCTATTAAA GAAAATGAGC TGCTGAAAAA AAAAAAAAAA		
PSP1-4	.....		
	2551		
PSP1-2	.		
PSP1-1	.		
PSP1-3	A		
PSP1-4	.		

FIGURE 2I/3

ページ( / / )

**FIGURE 3/3**

**Abstract**

Isolated nucleic acids encoding a human serine protease *PSP1*, protein obtainable from the nucleic acids, recombinant host cells transformed with the nucleic acids, oligonucleotides and primer pairs specific for *PSP1* polymorphisms and use of the protein and nucleic acid sequences are disclosed.